a./N 3

L SOCIETY

VLEDGE

L SOCIETY

Copyrighted material

Copyrighted material

Copyrighted material

Copyrighted material

a/n 3

PROCEEDINGS

OF THE

AMERICAN PHILOSOPHICAL SOCIETY

HELD AT PHILADELPHIA
FOR PROMOTING USEFUL KNOWLEDGE

VOLUME 84
1941

MDCCXXVII

THE AMERICAN PHILOSOPHICAL SOCIETY
PHILADELPHIA
1941

EDITORIAL BOARD

Editor-in-Chief	Edwin G. Conklin
Editor for Class I	Karl K. Darrow
Editor for Class II	Merkel H. Jacobs
Editor for Class III	Ernest M. Patterson
Editor for Class IV	Rhys Carpenter
Managing Editor	Arthur W. Goodspeed

LANCASTER PRESS, INC., LANCASTER, PA.

FROM A. P. S.
83106

CONTENTS

PAGE

Theoretical Synthesis of Supernova Spectra. FRED L. WHIPPLE AND CECILIA PAYNE-GAPOSCHKIN 1

Production of Secondary Electrons by Electrons of Energy Between 0.7 and 2.6 MEV. GEORGE HORNBECK AND IRL HOWELL 33

Autonomous Versus Reflexogenous Activity of the Central Nervous System. PAUL WEISS 53

Responsive Bone. C. B. DAVENPORT 65

Natural Selection Before the "Origin of Species." CONWAY ZIRKLE 71

Commemoration of the Life and Work of Alexander Dallas Bache.

Alexander Dallas Bache and His Connection with the American Philosophical Society. EDWIN G. CONKLIN 125

Alexander Dallas Bache and His Connection with The Franklin Institute of the State of Pennsylvania. HENRY BUTLER ALLEN 145

The Connection of Alexander Dallas Bache with the University of Pennsylvania. EDWARD P. CHEYNEY 151

Bache as an Educator. MERLE M. ODGERS 161

Alexander Dallas Bache as Superintendent of United States Coast Survey, 1843–1867. LEO OTIS COLBERT 173

Alexander Dallas Bache, a Founder, First President and Benefactor of the National Academy of Sciences. FRANK B. JEWETT 181

Symposium on Geomagnetism.

Terrestrial Electricity in Relation to Geomagnetism. O. H. GISH 187

The Magnetic Survey of the United States. N. H. HECK 205

The Significance of Fossil Magnetism. A. G. MCNISH 225

Geomagnetic Observatories and Instruments. H. E. MCCOMB 239

Magnetic Work at Sea. H. F. JOHNSTON 257

Geomagnetism: World-Wide and Cosmic Aspects with Especial Reference to Early Research in America. J. A. FLEMING 263

Aurora and Geomagnetism. C. W. GARTLEIN 299

PAGE
Contributions of Ionospheric Research to Geomagnetism. L. V. Berkner ... 309
Correlations of Short Wave Radio Transmission Across the Atlantic with Magnetic Conditions. H. E. Hallborg ... 323
Magnetism and Its Uses. Paul R. Heyl ... 339
S. Pietro in Vincoli and the Tripartite Transept in the Early Christian Basilica. Richard Krautheimer ... 353
John Bowring and the Poetry of the Slavs. Arthur Prudden Coleman ... 431
Symposium on Recent Advances in Psychology.
Coalescence of Neurology and Psychology. Karl S. Lashley ... 461
The Genesis of Behavior Form in Fetus and Infant. Arnold Gesell ... 471
On the Nature of Associations. Wolfgang Köhler ... 489
Mental Abilities. Edward L. Thorndike ... 503
Psychoanalysis and Scientific Method. Carney Landis 515
Psychology and Defense. Robert M. Yerkes ... 527
Motivation, Learning and Adjustment. Edward C. Tolman ... 543
Determination of Limb Darkening in Eclipsing Binaries from Color-Index Observations. Zdeněk Kopal ... 565
The Production of Neutrons by the Cosmic Radiation. S. A. Korff ... 589
Structure and Development of Centrifuged Eggs and Early Embryos of *Drosophila melanogaster*. Ruth B. Howland ... 605
Experimental Studies on the Reproductive Physiology of the Male Spring Peeper, *Hyla crucifer*. Roberts Rugh ... 617
What Columbus Saw on Landing in the West Indies. Leonardo Olschki ... 633
The Effects of Heat and Ultraviolet Light on Certain Physiological Properties of Yeast. Thomas F. Anderson and B. M. Duggar ... 661
The Corroded Bronze of Corinth. Earle R. Caley ... 689
Aboriginal Australian String Figures. Daniel Sutherland Davidson ... 763
Index ... 903

THEORETICAL SYNTHESIS OF SUPERNOVA SPECTRA

FRED L. WHIPPLE AND CECILIA PAYNE-GAPOSCHKIN

Harvard College Observatory

(*Read April 18, 1940*)

ABSTRACT

The supernovæ have already been shown to be temporary stars the intrinsic brightness of which at maximum light greatly exceeds that of any other known star, and is indeed of the same order as the brightness of an external galaxy. The spectra of these unique objects appear to consist of extremely wide bands (200± Angstroms) not previously identified either as absorption or emission lines or bands of known substances.

As a working hypothesis the authors have assumed that the spectra of supernovæ consist (principally) of permitted emission lines of astrophysically common elements superimposed upon a continuous background. The lines are further assumed to be broadened by a Doppler effect of expansion in the atmosphere, a process not unlike that generally assumed for the production of the spectra of ordinary novæ. A total line breadth corresponding to a relative expansion of 12,000 km/sec (measured along the line of sight) was adopted for the supernovæ, in conformity with the observed breadths of some of the prominent spectral features.

Evidence of various kinds shows that the effective temperature of the supernovæ near maximum light is relatively low (of the order of 10,000° K). The continuous spectrum appears to decrease in intensity with the time, and the temperature to rise within the first two hundred days after maximum light.

Because of the overlapping of greatly widened emission lines at the velocities considered, the ordinary methods of comparing laboratory and stellar spectra were not applicable. It was necessary to combine adopted spectra of various atoms by the double process of widening the lines in proportion to wave-length (with inverted-parabolic profiles) and then summing the intensities at small intervals of wave-length, with the inclusion of suitable background continua. These synthesized spectra were in a form for direct comparison with the intensity distributions of the observed spectra, as inferred from microphotometer tracings of the spectra of two supernovæ as published by R. Minkowski.

Since the whole value of the method depended upon an accurate knowledge of the relative intensities of the lines of each atom, all the intensities were calculated from modern atomic theory by means of the summation rules and the Einstein probability coefficients. The detailed method necessarily differed from atom to atom, and laboratory intensities were used as checks when theory was available, or in some cases directly.

Theoretical determinations were made of the relative intensities of the permitted lines of nineteen spectra, including hydrogen, helium, carbon, nitrogen, oxygen, sodium, calcium and iron in various stages of ionization. Complete tables of the transitions and the theoretical intensities of all of the lines studied are presented.

By combining the spectra at various levels of temperature (ionization) it was possible to duplicate many of the features of the supernova spectra. The peculiar red shift with time, observed only in the blue region of the spectrum, is now explained as a fortuitous effect arising from changes in the emission spectra with increasing temperature.

Hydrogen appears to be relatively less abundant than in the atmospheres of ordinary stars or novæ, while iron appears to be relatively more abundant than in stellar atmospheres. Some novæ, however, have shown very intense spectra of iron.

The results of the investigation are favorable to the supposition that the supernovæ are not essentially different from ordinary novæ in their atmospheric phenomena, except for the scale of the outburst and the consequent details of some of the physical processes.

THE purpose of the present investigation is to examine the possibility that the spectra of supernovæ consist essentially of a continuum on which the bright lines of common atoms are superposed. Before more speculative interpretations can be accepted it is necessary to exhaust the possibilities of analysis in terms of recognized transitions of known atoms (or molecules) in states of excitation that have been observed astrophysically. This analysis is restricted by the requirement that the implied physical processes must be internally consistent and must also be consistent with other processes that are generally believed to obtain in related phenomena. In a simple and direct fashion we are attempting to duplicate the observed spectral phenomena without recourse to new processes and hitherto unobserved conditions.

Details of the spectra of five supernovæ have been published: Z Centauri in N.G.C. 5253 by Pickering (1897); SN 1926 N.G.C. 4303 and SN 1936 N.G.C. 4273 by Humason (1936); SN 1937 N.G.C. 1003 by Minkowski (1939) and by Strohmeier (1938); and SN 1937 N.G.C. 4182 by Minkowski (1939), by Popper (1937) and by Strohmeier (1937). There has been no agreement as to the identification of the few features that can be distinguished in these spectra. They appear, at any given interval after maximum, to be substantially similar, if not identical, though of the more recently observed supernova in N.G.C. 5907 Humason and Minkowski (1940) state, "Its spectrum suggests important deviations from the normal pattern hitherto observed in supernovæ."

Although the spectrum of Z Centauri was at first described by Miss Cannon (1913) as resembling Class R, measures by Johnson (1936) showed features incompatible with this classification, and confirmed the suggestions of Ritchey and Pease (1917), of Vorontsov-Velyaminov (1931) and of C. Payne-Gaposchkin (1936), that supernovæ have bright-line spectra. It is obvious, if this interpretation be correct, that the bright lines must be exceedingly wide.

Recent investigations, notably that of Minkowski (1939), show that the spectra of supernovæ undergo striking changes as a function of time after maximum light. In the blue region of the spectrum the changes appear as a progressive shifting of the main features toward the red, an effect not shared by the features in the green, yellow, and red. A physical red-shift that affects only a limited region of the spectrum is difficult to envisage. In the present investigation it will be shown that this shift can be explained without recourse to special physical processes.

Our method is to assume that emission lines are enormously widened and that their profiles combine additively in the formation of the spectrum, which is superposed upon a continuum. For this purpose it is necessary to decide upon the appropriate atomic spectra, to determine the relative intensities of the lines, and to choose a suitable continuum. In the choice of spectra to be used we are guided by astrophysical abundance, and limited by laboratory information and theory based upon it. All possible atomic spectra of hydrogen, helium, carbon, nitrogen, and oxygen, all available spectra of iron, the spectrum of neutral sodium and of once ionized calcium have been used in the synthesis. This list includes all the commoner atomic species except magnesium and silicon, which, upon investigation, did not appear to contribute appreciably. In the following section are presented the methods of obtaining the relative intensities of the lines in these spectra, and the numerical values.

THE EMISSION LINE INTENSITIES

The relative intensities of the emission lines in the spectrum of a given atom in a particular stage of ionization are given by the formula:

$$I \propto \pi_n \nu A_{nn'} \exp\left(-\chi_n/kT\right), \tag{1}$$

where n and n' are the respective quantum numbers of the lower and upper states, π is the statistical weight, and $A_{nn'}$ is the Einstein probability factor. Since numerical values of the A's were available only for H I, He I, and He II, it was necessary to make estimates of the transition probabilities for the other elements. Within multiplets it was possible to use the strengths given by Russell (1936), and for the relative strengths of multiplets in transition arrays, the tables given by Goldberg (1935). The strengths, S, are related to the A's by the following relation:

$$A_{nn'} \propto \frac{1}{\pi_n} S_{nn'} \nu^3. \tag{2}$$

For certain spectra it was possible to estimate values of the Ladenburg factor, f, by means of the sum rule (Kuhn and Thomas-Reiche, 1938). Line intensities based on these estimates are correct only in order of magnitude. This procedure is justified for relatively simple atoms. The A's may be calculated from the f's by the relation:

$$A_{nn'} \propto \frac{\pi_{n'}}{\pi_n} f_{n'n} \nu^2. \tag{3}$$

For the hydrogenic atoms, C IV, N V, and O VI, the radial quantum integrals, $c\sigma^2$, as derived by Condon and Shortley (1935) were utilized. The relevant transitions are those between levels of high orbital quantum number so that the assumption of hydrogenic wave functions is justified.

For certain complicated spectra, notably Fe II and Fe III, the assumption of Russell-Saunders coupling is inapplicable, except within certain multiplets, so that the methods of the preceding paragraph cannot be used. The only procedure available was the adoption of intensities on the basis of laboratory and astrophysical determinations.

For all spectra where laboratory intensities were available, theoretical intensities were checked by comparison of the two. The relation published by Russell (1925) showed that the squares of the laboratory intensities determined by A. S. King are proportional to the true intensities. This proportionality was generalized to apply to all laboratory intensities, and was found to be statistically quite accurate, when a suitable temperature was adopted in formula (1).

The detailed method of obtaining the theoretical intensities for each of the atomic spectra is described in a brief paragraph preceding the corresponding table below. These methods are summarized in Table I; the fundamental quantities adopted for the calculation are listed in the column headed "Method."

TABLE I

Method	Atomic Spectra
A	H I, He I, He II
S, I_{lab}	N II, O II, Fe II
S, f	C III, N IV, O III, O IV, O V, Na I, Ca II
S, f, I_{lab}	C II, N III
$c\sigma^2$	C IV, N V, O VI
I_{lab}	Fe III

For each atomic spectrum an appropriate temperature was assumed, ranging from 15,000° K for hydrogen to 100,000° K for O VI: these temperatures are given in the descriptive paragraphs. In some of the spectra a very large range in excitation potential is represented. Those lines that would be relatively strengthened at higher temperatures (high E.P.) are marked in the tables with double daggers; those that would be relatively strengthened at lower temperatures (low E.P.), with daggers. Because of the great widths assumed for the lines in the synthesized spectra, multiplets not covering a great range in wave-length have been summed in intensity and treated as a single line. If the intensities of the individual lines are desired, they can be calculated from the tabulations of the strengths of lines within multiplets as given by Russell (1925). The headings of the tables are self-explanatory; our adopted relative intensity, *I*, is given in the last column of each. The strongest line (or multiplet) is generally taken to have an intensity 100. The wave-lengths for multiplets are means weighted according to the intensities.

Hydrogen I—The intensities were calculated by means of the formula:

$$I \propto n^2 \nu A_{nn'} \exp\left(-\chi_n/kT\right), \tag{1a}$$

where n' and n are the respective quantum numbers of the lower and upper states. The values of the Einstein probability coefficients $A_{nn'}$ were taken from Table IIa of the theoretical discussion by Menzel and Pekeris (1935). The adopted temperature was 15,000° K. Since the hydrogen lines clearly do not attain any great intensity in the spectra of supernovæ it was considered unnecessary to calculate the effect of the Balmer continuum, either in emission or in absorption.

TABLE 2

SPECTRUM OF HYDROGEN

λ	Combination	*I*
3798	$2^2P^\circ - 10^2D$ etc.	4
3835	$2^2P^\circ - 9^2D$ "	6
3889	$2^2P^\circ - 8^2D$ "	8
3970	$2^2P^\circ - 7^2D$ "	13
4102	$2^2P^\circ - 6^2D$ "	21
4340	$2^2P^\circ - 5^2D$ "	41
4861	$2^2P^\circ - 4^2D$ "	100
6563	$2^2P^\circ - 3^2D$ "	362

Helium I—The intensities were calculated by means of the formula:

$$I \propto \pi_n \nu A_{nn'} \exp\left(-\chi_n/kT\right). \tag{1}$$

The weights π_n were calculated from the relation:

$$\pi_n = (2S + 1)(2L + 1).$$

The values of $A_{nn'}$ were taken from the calculations by Goldberg

(1939), except for the line at 3187, for which the value of $A_{nn'}$ was calculated to be 5.54×10^6. The adopted temperature was 15,000° K. The laboratory intensities are taken from C. E. Moore (1933).

TABLE 3

SPECTRUM OF HELIUM I

λ	Combination	I_{lab}	I
3188	$2^3S - 4^3P^\circ$	8	6
3555	$2^3P^\circ - 10^3D$	1,1	1
3587	$2^3P^\circ - 9^3D$	2,1	2
3634	$2^3P^\circ - 8^3D$	2,1	3
3705	$2^3P^\circ - 7^3D$	3,1	4
3820	$2^3P^\circ - 6^3D$	4,1	7
3889	$2^3S - 3^3P^\circ$	10	16
3927	$2^1P^\circ - 8^1D$	1	1
3965	$2^1S - 4^1P^\circ$	4	2
4009	$2^1P^\circ - 7^1D$	1	1
4026	$2^3P^\circ - 5^3D$	5,1	14
4144	$2^1P^\circ - 6^1D$	2	1
4388	$2^1P^\circ - 5^1D$	3	3
4472	$2^3P^\circ - 4^3D$	6,1	32
4922	$2^1P^\circ - 4^1D$	4	7
5016	$2^1S - 3^1P^\circ$	6	5
5876	$2^3P^\circ - 3^3D$	10,1	100
6678	$2^1P^\circ - 3^1D$	6	31

Helium II—The intensities were calculated as for hydrogen, the values of $A_{nn'}$ being taken from Menzel and Pekeris (1935). The adopted temperature was 30,000° K. As a result of the difference in excitation, the lines at 4686 and 3203 would be somewhat intensified at lower temperatures, relative to those of the Pickering series.

TABLE 4

SPECTRUM OF HELIUM II

λ	Combination	I
3203	$3^2D - 5^2F^\circ$ etc.	33†
4100	$4^2F^\circ - 12^2G$ "	1
4200	$4^2F^\circ - 11^2G$ "	1
4339	$4^2F^\circ - 10^2G$ "	1
4542	$4^2F^\circ - 9^2G$ "	2
4686	$3^2D - 4^2F^\circ$ "	100†
4859	$4^2F^\circ - 8^2G$ "	2
5412	$4^2F^\circ - 7^2G$ "	4
6560	$4^2F^\circ - 6^2G$ "	7

Carbon II—The analysis, notation, and laboratory intensities were taken from Edlén (1933—1). The intensities were calculated by means of the formula:

$$I \propto \pi_{n'} f \nu^3 \exp\,(-\,\chi_n/kT), \tag{4}$$

f being the oscillator strength. A temperature of 25,000° K was used. Values of f were estimated from the character of the Grotrian diagram. The weights $\pi_{n'}$ were made mutually consistent within the primed and unprimed systems respectively. The primed system arises from a parent configuration $1s^2 2s 2p$, the unprimed system, from a parent configuration $1s^2 2s^2$. The relative intensities for the two systems were adjusted by means of the laboratory intensities of all the lines included in the table, on the assumption that the squares of the laboratory intensities were proportional to the true intensities of the lines. The table contains summed intensities of whole multiplets, the relative intensities of which were deduced by means of the tables given by Goldberg (1935).

TABLE 5

SPECTRUM OF CARBON II

λ	Combination	I_{lab}	f	I
3361	$3d\,{}^2D - 5p\,{}^2P$	2,3	0.1	18
3585	$3p'\,{}^4D - 4s'\,{}^4P$	2,4,1,2,2,1	0.3	23
3877	$3d'\,{}^4F - 4f'\,{}^4G$	6,8,7,6,1,1,1	1.2	42
3920	$3p\,{}^2P - 4s\,{}^2S$	10,9	0.4	74
3949	$3d'\,{}^4F - 4f'\,{}^4F$	2,1,0	1.2	4
3977	$3d'\,{}^4D - 4f'\,{}^4D$	3,1	1.2	3
4015	$3p'\,{}^4S - 4s'\,{}^4P$	2,1,0	0.3	3
4075	$3d'\,{}^4D - 4f'\,{}^4F$	7,5,6	1.2	24
4267	$3d\,{}^2D - 4f\,{}^2F$	20,19	0.8	100
4290	$3d'\,{}^2D - 4f'\,{}^2D$	1,1	1.2	1
4318	$3p'\,{}^4P - 4s'\,{}^4P$	2,0,2,4,2	0.3	6
4372	$3d'\,{}^4P - 4f'\,{}^4D$	5,4,3,4	1.2	12
4411	$3d'\,{}^2D - 4f'\,{}^2F$	5,5	1.2	9
4619	$3d'\,{}^2F - 4f'\,{}^2G$	5	1.2	10
4730	$3d'\,{}^2F - 4f'\,{}^2F$	2,1	1.2	1
4960	$3p'\,{}^2P - 3d'\,{}^2P$	2,0,1	1.0	2
5046	$3d'\,{}^4D - 4p'\,{}^4P$	0,1	0.5	5
5116	$3d'\,{}^2P - 4f'\,{}^2D$	2,2	1.2	3
5140	$3s'\,{}^4P - 3p'\,{}^4P$	3,2,0,1,2,5,3	1.6	35
5337	$4p\,{}^2P - 6s\,{}^2S$	0	0.2	3
5536	$4s\,{}^2S - 5p\,{}^2P$	1	1.8	15
5650	$3s'\,{}^4P - 3p'\,{}^4S$	2,3,4	1.6	9
5890	$3d\,{}^2D - 4p\,{}^2P$	3,4	0.6	41
6098	$3p'\,{}^2P - 3d'\,{}^2D$	0,3,2	1.0	4
6115	$3s'\,{}^2P - 3p'\,{}^2S$	0	1.6	1
6580	$3s\,{}^2S - 3p\,{}^2P$	9,10	1.1	63
6780	$3s'\,{}^4P - 3p'\,{}^4D$	4,2,6,2,3,1,3,0	1.6	25

Carbon III—The analysis, notation, and laboratory intensities were taken from Edlén (1933—1 and 1933—2). The intensities were calculated as for Carbon II, except that the weights for the primed system (parent configuration $1s^22p$) were increased by a factor of 3 relative to the unprimed system (parent configuration $1s^22s$). The correctness of this assumption was verified by means of the laboratory intensities.

A temperature of 40,000° K was used. The lines of relatively high excitation potential (strengthened at higher temperatures) are marked in the table by double daggers, those of relatively low excitation potential, with daggers.

TABLE 6

SPECTRUM OF CARBON III

λ	Combination	I_{lab}	f	I
3609	$4p^3P - 5d^3D$	5,4	0.2	5
3704	$3p'^1P - 3d'^1P$	2	1.0	0‡
3887	$4d^3D - 5f^3F$	2,3,4	0.12	4
4000	$4d'^3F - 5f'^3G$?	0	0.12	1‡
4056	$4d^1D - 5f^1F$	5	0.12	1
4069	$4f^3F - 5g^3G$	9,10,10	1.6	62
4122	$4p^1P - 5d^1D$	3	0.2	1
4187	$4f^1F - 5g^1G$	10	1.6	19
4326	$3s'^1P - 3p'^1D$	8	1.2	6
4362	$4f'^1G - 5g'^1H$?	2	1.6	2‡
4368	$4f'^3G - 5g'^3H$?	4	1.6	6‡
4383	$4d^3D - 5p^3P$	0,1,2	0.5	12
4516	$4p^3P - 5s^3S$	4,3	0.25	3
4594	$4f'^3F - 5g'^3G$?	2	1.6	4‡
4650	$3s^3S - 3p^3P$	20,19,18	1.2	100†
4668	$3s'^3P - 3p'^3P$	4,6,4	1.2	9
5250	$4d^1D - 5p^1P$	0	0.5	2
5255	$3s'^3P - 3p'^3S$	0,1,2	1.2	2
5696	$3p^1P - 3d^1D$	8	1.0	25†
5872	$3p'^3D - 3d'^3P$	—	1.0	0
6738	$3s'^3P - 3p'^3D$	—	1.2	4

Carbon IV—The analysis and notation are those of Edlén (1933—2). The quantity $c\sigma^2$ is equal to

$$(4l^2 - 1)^{-1}\left[\int_0^\infty rR(nl)R(n'l - 1)dr\right]^2,$$

the radial quantum integral in atomic units (Condon and Shortley, 1935). Values of the radial quantum integrals for transitions not included by Condon and Shortley were extrapolated graphically by the writers; such values are given to two decimal places only, in the third column of the table.

The intensities were calculated by means of the formula:

$$I \propto (c\sigma^2)\nu^4 \exp\,(-\ \chi_n/kT). \tag{5}$$

A temperature of 100,000° K was used. The lines at λ5805, which are nearly twenty volts lower in excitation than the other lines, are greatly intensified relatively at lower temperatures.

TABLE 7

SPECTRUM OF CARBON IV

λ	Combination	$\log c\sigma^2$	I
3934	$5S-6P$	0.951	45
4441	$5P-6D$	0.764	18
4646	$5D-6F$	0.62	11
4656	$5F-6G$	0.52	8
4659	$5G-6H$	0.28	5
4664	$5F-6D$	9.64	1
4685	$6G-8H$	9.84	1
4685	$6H-8I$	9.71	1
4786	$5D-6P$	0.06	3
5017	$5P-6S$	0.54	7
5805	$3S-3P$	1.034	100†
6592	$6S-7P$	1.42	14

Nitrogen II—The analysis, notation, and laboratory intensities are taken from Edlén (1933–1) and from Freeman (1929). Within each supermultiplet the intensities of multiplets were calculated with the aid of the tables given by Goldberg (1935). A temperature of 25,000° K was used. The total intensities of the supermultiplets were adjusted by means of the laboratory intensities, on the assumption that the squares of the laboratory intensities are proportional to the true intensities.

On the assumption that the laboratory temperature also corresponds to 25,000° K, the above calculations lead to the following values of f for various transitions. To allow for the change in the parent configuration ($1s^22s^22p$ for singlets and triplets; $1s^22s2p^2$ for quintets) the weight factors were multiplied by three for the quintets.

Singlets and Triplets		*Quintets*	
Transition	f	Transition	f
$3s-3p$	1.1	$3s-3p$	1.1
$3p-4s$	0.2		
$3p-3d$	1.0	$3p-3d$	1.0
$3d-4p$	0.4		
$3d-4f$	0.4		

TABLE 8

SPECTRUM OF NITROGEN II

λ	Combination	I_{lab}	I
3007	$3p^1P-4s^1P$	7	16
3329	$3p^3D-4s^3P$	2,2,4,2,3	53
3437	$3s^1P-3p^1S$	6	17
3609	$3p^3S-4s^3P$	3,2,1	8
3845	$3p^3P-4s^3P$	3,5,3,3,2,3	18
3919	$3p^1P-3d^1P$	6	13
3995	$3s^1P-3p^1D$	10	74
4038	$3d^3F-4f^3G$	5,4	30
4045	$3d^1D-4f^1D$	1	1
4057	$3d^1D-4f^1F?$	1	7
4078	$3d^3F-4f^3F$	2,2	3
4134	$3s^5P-3p^5S$	1,2,3	8
4171	$3d^3D-4f^3D$	0,0,1	2
4228	$3p^1D-4s^1P$	3	7
4240	$3d^3D-4f^3F$	8,6	17
4434	$3d^3P-4f^3D$	2,0,6,2,3	10
4447	$3p^1P-3d^1D$	10	28
4465	$3d^1F-4f^1F$	—	1
4485	$3p^3D-3d^3P$	1,0,2,0,3	2
4530	$3d^1F-4f^1G?$	5	6
4610	$3d^1P-4f^1D$	0	3
4625	$3s^3P-3p^3P$	8,7,6,7,10,8	90
4709	$3p^5D-3d^5D$	1,0,0,1,0,2	4
4792	$3p^3D-3d^3D$	2,4,2,5,2,6,2	20
4860	$3p^5D-3d^5P$	2	0
5000	$3p^3S-3d^3P$	4,3,7	21
5005	$3s^5P-3p^5P$	2,3,0,5,1,2,2	14
5015	$3p^3D-3d^3F$	7,8,5,5,6,0	100
5030	$3s^3P-3p^3S$	2,6,8	24
5104	$3p^1S-4s^1P$	2	1
5177	$3p^5P-3d^5D$	0,2,2,1,1,1,1	9
5180	$3p^5D-3d^5F$	1,2,3,5,2	18
5340	$3p^5P-3d^5P$	3,2,1,4	3
5477	$3d^1D-4p^1D$	—	3
5480	$3p^3P-3d^3P$	2,2,3,2,3,5	11
5538	$3s^5P-3p^5D$	2,4,5,1,3,3,3,00,0	17
5686	$3s^3P-3p^3D$	8,6,10,6,6,2	83
5940	$3p^3P-3d^3D$	4,7,2,8,3,0	26
6168	$3d^1P-4p^1S$	4	2
6170	$3d^3F-4p^3D$	4,1,3	33
6243	$3d^1F-4p^1D$	5	11
6284	$3p^1D-3d^1P$	3	1
6347	$3d^3D-4p^3P$	1,4,1,3	16
6482	$3s^1P-3p^1P$	8	11
6520	$3d^3D-4p^3D$	0,1,0,0,2	5
6566	$3d^1P-4p^1D$	—	0
6611	$3p^1D-3d^1F$	6	9
6630	$3d^1D-4p^1P$	2	5
6870	$3p^5S-3d^5P$	2,1,1	1

Nitrogen III—The analysis and notation are those of Edlén (1933—1). The laboratory intensities for the doublet system were taken from Edlén (1933—1); those for the quartet system, from Freeman (1928). All of the intensities were derived theoretically by the method used for Carbon II, except for the supermultiplet $2s2p3p' - 2s2p3d'$, for which the assumed value $f = 1.0$ was adjusted to $f = 1.2$ by means of the laboratory intensities. The weights for the primed transitions (parent configuration $1s^22s2p$) were increased by a factor of three as compared to the weights for the unprimed transitions (parent configuration $1s^22s^2$). A temperature of 40,000° K was used. The lines at 4638 and 4100 would be relatively strengthened at lower temperatures because their excitation potential is about ten volts lower than those of the other lines.

TABLE 9

Spectrum of Nitrogen III

λ	Combination	I_{lab}	f	I
3350	$3s'\,{}^2P - 3p'\,{}^2S$	2,1	1.0	5
3360	$3s'\,{}^4P - 3p'\,{}^4P$	4,4,1,2,3,7,6	1.0	58
3750	$3s'\,{}^4P - 3p'\,{}^4S$	4,6,7	1.0	15
3792	$3p'\,{}^4D - 3d'\,{}^4P$	1	1.2	1
3938	$3p'\,{}^2P - 3d'\,{}^2D$	3,4,1	1.2	18
4001	$4d\,{}^2D - 5f\,{}^2F$	3,4	0.1	2
4100	$3s\,{}^2S - 3p\,{}^2P$	9,10	0.8	84†
4205	$3s'\,{}^2P - 3p'\,{}^2D$	3,6,5	1.0	21
4335	$3p'\,{}^4D - 3d'\,{}^4D$	2,5,3,4,2,2,3,2,1	1.2	10
4379	$4f\,{}^2F - 5g\,{}^2G$	10	2.2	38
4525	$3s'\,{}^4P - 3p'\,{}^4D$	6,7,3,4,1,3,0	1.0	52
4536	$3p'\,{}^4S - 3d'\,{}^4P$	3,2,0	1.2	10
4545	$4p\,{}^2P - 5s\,{}^2S$	0	0.1	1
4638	$3p\,{}^2P - 3d\,{}^2D$	7,10,9	1.0	100†
4865	$3p'\,{}^4D - 3d'\,{}^4F$	0,1,0,2,5,4,3,2	1.2	43
5290	$3p'\,{}^4P - 3d'\,{}^4P$	2,1,1,00,1,1,1	1.2	5
6466	$3p'\,{}^4P - 3d'\,{}^4D$	0,2,00,4,2,3,2,2	1.2	9

Nitrogen IV—The analysis and notation are taken from Edlén (1933—1 and 1933—2). All intensities were derived theoretically by the method used for Carbon II. The weights for the primed transitions (parent configuration $1s^22p$) were increased by a factor of three relative to those for the unprimed transitions (parent configuration $1s^22s$). The adopted temperature was 100,000° K. At lower temperatures the lines at λλ3480, 4058 and 6383 would be relatively intensified because their excitation potentials are some ten volts lower than the average level.

TABLE 10

SPECTRUM OF NITROGEN IV

λ	Combination	f	I
3463	$3s'{}^3P - 3p'{}^3P$	1.2	3
3480	$3s^3S - 3p^3P$	1.2	100†
3700	$3p'{}^3D - 3d'{}^3P$	1.0	1
3748	$3s'{}^1P - 3p'{}^1D$	1.2	12
4058	$3p^1P - 3d^1D$	1.0	37†
4178	$3p'{}^3S - 3d'{}^3P$	1.0	11
4505	$3s'{}^3P - 3p'{}^3S$	1.2	0
4743	$3p'{}^3D - 3d'{}^3D$	1.0	6
5073	$3p'{}^1D - 3d'{}^1P$	1.0	0
5240	$3s'{}^3P - 3p'{}^3D$	1.2	1
5734	$3p'{}^1P - 3d'{}^1D$	1.0	4
5830	$3p'{}^3P - 3d'{}^3P$	1.0	3
6383	$3s^1S - 3p^1P$	1.2	5†

Nitrogen V—The analysis and notation are taken from Edlén (1933—1 and 1933—2). The calculations were made exactly as for Carbon IV. The line at λ4609 ($3S - 3P$) has an excitation more than thirty volts lower than those of the other lines, and would be intensified at temperatures lower than the assumed value of 100,000° K.

TABLE 11

SPECTRUM OF NITROGEN V

λ	Combination	$\log c\sigma^2$	I
3161	$5P - 6S$	0.54	64
4335	$6S - 7P$	1.42	100
4609	$3S - 3P$	1.034	1240†
4751	$6P - 7D$	0.89	20
4933	$6D - 7F$	0.90	18
4943	$6F - 7G$	0.54	8
4945	$6G - 7H$	0.45	6
4945	$6H - 7I$	0.11	3
4952	$6F - 7D$	0.28	4
5067	$6D - 7P$	0.62	8
5273	$6P - 7S$	0.92	15
6719	$7S - 8P$	1.72	29

Oxygen II—The analysis, notation, and intensities were taken from Edlén (1933—1 and 1935) when possible. For other lines, not specifically tabulated by Edlén, the intensities were taken from Fowler (1926). Relative intensities of multiplets within the three transition arrays $3s - 3p$; $3p - 3d$; and $3d - 4f$ were calculated with the aid of Goldberg's tables, as was done for Nitrogen II. The relative intensities of the transition arrays were adjusted by means of the laboratory intensities. A temperature of 30,000° K was adopted. The total range in level between the lines of low and high excitation is about eight volts. The intensities marked with colons were deduced from the laboratory intensities alone.

TABLE 12

SPECTRUM OF OXYGEN II

λ	Combination	I_{lab}	I
3130	$3p\,{}^4D^\circ - 4s\,{}^4P$	1,6,2,7,3,10,8,4	84
3272	$\overline{3p}\,{}^2F^\circ - \overline{4s}\,{}^2D$	7,7	18
3290	$3p\,{}^4P^\circ - 4s\,{}^4P$	7,9,5,4,3,6,6	41
3384	$3p\,{}^2S^\circ - 3d\,{}^2P$	7,8	23
3408	$\overline{3p}\,{}^2D^\circ - \overline{4s}\,{}^2D$	7,6	11
3465	$3p\,{}^2D^\circ - 4s\,{}^2P$	5,5,8	25
3720	$3s\,{}^4P - 3p\,{}^4S^\circ$	7,8,9	33†
3732	$\overline{3p}\,{}^2P^\circ - 4s\,{}^2D$	2,3	5
3750	$3p\,{}^4S^\circ - 4s\,{}^4P$	6,5,4	8
3805	$3p\,{}^2P^\circ - 4s\,{}^2P$	3,6,4,4	11
3856	$3p\,{}^4D^\circ - 3d\,{}^4D$	3,3,3,3,5,4,2,1,4,7	21
3900	$3p\,{}^4D^\circ - 3d\,{}^4P$	1,2,1,2,1,4	1
3913	$\overline{3s}\,{}^2D - \overline{3p}\,{}^2P^\circ$	10,2,6	16
3971	$3s\,{}^2P - 3p\,{}^2P^\circ$	5,7,10,5	37†
4024	$\overline{3d}\,{}^2F - \overline{4f}\,{}^2D^\circ$	1	0‡
4048	$3d\,{}^4F - 4f\,{}^4F^\circ$	3,2,00,0,0,0,0,0	1
4054	$\overline{3d}\,{}^2F - \overline{4f}\,{}^2F^\circ$	00,0	1‡
4061		2,3	3:
4080	$3p\,{}^4D^\circ - 3d\,{}^4F$	4,6,8,10,4,3,5,0,0	100
4087	$\overline{3d}\,{}^2F - \overline{4f}\,{}^2G^\circ$	4	3‡
4095	$3d\,{}^4F - 4f\,{}^4G^\circ$	0,6,0,3,7	16
4107	$3d\,{}^4F - 4f\,{}^4D^\circ$	1	0
4110	$3p\,{}^4P^\circ - 3d\,{}^4D$	3,5,4,7,3,4,8,2	48
4112	$\overline{3p}\,{}^2F^\circ - \overline{3d}\,{}^2D$	1,1	1
4143	$3p\,{}^4P^\circ - 3d\,{}^4P$	4,2,6,0,7,3,4	16
4144	$3p'\,{}^6P - 3d'\,{}^6D$	4,0,2,1,0,1,0	5:
4187	$\overline{3p}\,{}^2F^\circ - \overline{3d}\,{}^2G$	8,10	21
4194	$\overline{3p}\,{}^2D^\circ - \overline{3d}\,{}^2P$	2,00,1	2
4253	$\overline{3d}\,{}^2G - \overline{4f}\,{}^2H^\circ$	4,4	3‡
4280	$3d\,{}^4D - 4f\,{}^4F^\circ$	1,0,2,1,1,3,0,4	10
4300	$3d\,{}^4P - 4f\,{}^4D^\circ$	1,0,4,3,1,0	6
4314	$3d\,{}^2F - 4f\,{}^2F^\circ$	0,0	1
4324	$\overline{3p}\,{}^2P^\circ - \overline{3d}\,{}^2S$	1,2	2
4333	$\overline{3d}\,{}^2G - \overline{4f}\,{}^2G^\circ$	0	0‡
4330	$\overline{3p}\,{}^2D^\circ - \overline{3d}\,{}^2D$	3,0,0,2	6
4336	$3s\,{}^4P - 3p\,{}^4P^\circ$	8,8,3,6,7,8,7	68†
4341	$3d\,{}^2F - 4f\,{}^2G^\circ$	3,1	6
4344	$\overline{3d}\,{}^2D - \overline{4f}\,{}^2D^\circ$	1,1	1‡
4349	$\overline{3s}\,{}^2D - \overline{3p}\,{}^2D^\circ$	5,6	20

TABLE 12—*Continued*

λ	Combination	I_{lab}	I
4350	$3d{}^4D - 4f{}^4D^\circ$	1,1,0	1
4378	$\overline{3d}{}^2D - \overline{4f}{}^2F^\circ$	0,0	1‡
4390	$3p{}^2D^\circ - 3d{}^2D$	1,4,7,1	6
4425	$3s{}^2P - 3p{}^2D^\circ$	8,10,6	45†
4446	$\overline{3p}{}^2F^\circ - \overline{3d}{}^2F$	5,6	5
4465	$3s'{}^6S - 3p'{}^6P$	4,4,1,4,1	10:
4488	$\overline{3d}{}^2P - \overline{4f}{}^2D^\circ$	2,0	1‡
4490	$3d{}^2P - 4f{}^2D^\circ$	3,2	3
4592	$\overline{3s}{}^2D - \overline{3p}{}^2F^\circ$	9,8	24
4608	$3d{}^2D - 4f{}^2F^\circ$	0,5,1	4
4650	$3s{}^4P - 3p{}^4D^\circ$	6,9,10,6,9	92†
4678		4,8,2	
4698	$\overline{3p}{}^2P^\circ - \overline{3d}{}^2P$	0,1,2,0	3
4700	$\overline{3p}{}^2D^\circ - \overline{3d}{}^2F$	1,7,3	7
4708	$3d{}^2D - 4f{}^2D^\circ$	0	0
4710	$3p{}^2D^\circ - 3d{}^2F$	8,5,2	26
4843	$\overline{3d}{}^2S - \overline{4f}{}^2P^\circ$?	0	0‡
4868	$\overline{3p}{}^2P^\circ - \overline{3d}{}^2D$	3,5	2
4910	$3p{}^4S^\circ - 3d{}^4P$	4,5,6	11
4945	$3p{}^2P^\circ - 3d{}^2D$	5,7,3	11
5185	$3p{}^2P^\circ - 3d{}^2P$	4,2,3,5	3
6680	$3s{}^2P - 3p{}^2S^\circ$	4,5	2†

Oxygen III—The analysis and notation are from Edlén (1933—1). The intensities are from Miss Moore (1933) for the lines that she tabulates, and from Edlén (1933—1) for other lines. All the intensities were derived theoretically by the method used for Carbon II. The weight factors for transitions of parent configuration $1s^22s2p^2$ were taken to be three times those for $1s^22s^22p$. Instead of the factor 3, a factor of 4.4 would be derived by comparison with the laboratory intensities. Since the transitions of $1s^22s^22p$ are of lower excitation by about ten volts, the discrepancy would be removed by the adoption of a temperature of 46,000° K, instead of the temperature of 40,000° K actually used. The theoretical values of the intensities were therefore retained.

Oxygen IV—Analysis, notation, and intensities are taken from Edlén (1933—1 and 1933—2). The calculations were carried out as for Carbon II, with an adopted temperature of 100,000° K. The value of f was 1.0 for all lines excepting λ3067, for which a value of 0.8 was used, and λ3836, for which 0.1 was used. The statistical weights for transitions of parent term ^3P were taken to be three times those for transitions of parent term ^1P and ^1S.

TABLE 13

SPECTRUM OF OXYGEN III

λ	Combination	I_{lab}	f	I
3008	$3p^3D - 3d^3D^{\circ}$	3,4,5	1.0	21†
3040	$3s^3P^{\circ} - 3p^3P$	5,4,4,5,8,6	1.1	76†
3080	$3p'^5D^{\circ} - 3d'^5D$	2,1	1.0	4
3122	$3p^3S - 3d^3P^{\circ}$	4,5,6	1.0	27†
3204	$3p'^3D^{\circ} - 3d'^3D$	1,1	1.0	2
3265	$3p^3D - 3d^3F^{\circ}$	8,10,5	1.0	100†
3315	$3s^3P^{\circ} - 3p^3S$	3,5,6	1.1	22†
3340	$3p'^5P^{\circ} - 3d'^5P$	3,2,2,3,1,4,0	1.0	3
3350	$3s'^5P - 3p'^5P^{\circ}$	4,4,4,4,3	1.1	13
3380	$3p'^5P^{\circ} - 3d'^5D$	1,2,3,4	1.0	9
3436	$3p^3P - 3d^3P^{\circ}$	3,4,5	1.0	13†
3450	$3p'^5D^{\circ} - 3d'^5F$	2,4,2,5,2,2,1	1.0	17
3557	$3s'^3P - 3p'^3P^{\circ}$	1	1.1	4
3644	$3p'^3P^{\circ} - 3d'^3D$	1,0,00,2,1,3	1.0	3
3700	$3s'^5P - 3p'^5D^{\circ}$	5,5,4,3,3,2	1.1	18
3708	$3p^3P - 3d^3D^{\circ}$	5,6,6	1.0	33†
3729	$3p'^3D^{\circ} - 3d'^3F$	1,1,0	1.0	6
3760	$3s^3P^{\circ} - 3p^3D$	7,5,9,6,6,2	1.1	86†
3962	$3s^1P^{\circ} - 3p^1D$	8	1.1	20†
4077	$3s'^3P - 3p'^3D^{\circ}$	1,0	1.1	6
4450	$3p'^5S^{\circ} - 3d'^5P$	1,0	1.0	2
4555	$3p'^3P^{\circ} - 3d'^3P$	0	1.0	1
5268	$3p^1S - 3d^1P^{\circ}$	2	1.0	2†
5592	$3s^1P^{\circ} - 3p^1P$	6	1.1	6†

TABLE 14

SPECTRUM OF OXYGEN IV

λ	Combination	I_{lab}	I
3042	$3s'^2P - 3p'^2S$	1,0	3
3067	$3s^2S - 3p^2P$	5,6	56†
3190	$3p'^4D - 3d'^4D$	—	6
3350	$3s'^2P - 3p'^2D$	0,3,2	7
3360	$3p'^4S - 3d'^4P$	0,1,2	7
3386	$3s'^4P - 3p'^4D$	4,5,2,2	19†
3411	$3p^2P - 3d^2D$	4,5,1	100†
3491	$\overline{3s}'^2P - \overline{3p}'^2D$	1,0	5‡
3512	$3p'^2S - 3d'^2P$	—	3
3562	$3p'^2D - 3d'^2F$	1,2	11
3730	$3p'^4D - 3d'^4F$	3,3,4	25
3836	$5d^2D - 6f^2F$	—	1‡
3840	$\overline{3p}'^2P - \overline{3d}'^2D$	—	1‡
3982	$3p'^4P - 3d'^4P$	0,1	3
4262	$\overline{3p}'^2D - \overline{3d}'^2F$	—	1‡
4472	$4p'^2P - 4d'^2P$	—	0
4570	$5f^2F - 6d^2D$	—	11‡
4791	$3p'^4P - 3d'^4D$	2,3	6
5334	$3p'^2D - 3d'^2D$	—	1
5440	$4p'^2P - 4d'^2D$	—	1‡

Oxygen V—Analysis and notation are taken from Edlén (1933—1 and 1933—2). The calculations were carried out as for Carbon II, with an adopted temperature of 100,000° K. The statistical weights for the transitions arising from primed terms were taken to be three times those for transitions from unprimed terms.

TABLE 15

SPECTRUM OF OXYGEN V

λ	Combination	f	I
3059	$3s'{}^1P - 3p'{}^1D$	1.2	34
3276	$3s'{}^3P - 3p'{}^3S$	1.2	20
3732	$3p'{}^3D - 3d'{}^3D$	1.0	19
4124	$3s'{}^3P - 3p'{}^3D$	1.2	55
4147	$3p'{}^3S - 3d'{}^3P$	1.0	17
4522	$3p'{}^1D - 3d'{}^1P$	1.0	0
4554	$3p'{}^1P - 3d'{}^1D$	1.0	12
5114	$3s^1S - 3p^1P$	1.2	23†
5450	$3p'{}^3P - 3d'{}^3P$	1.0	6
5590	$3p^3P - 3d^3D$	1.0	100†
6329	$3p'{}^1D - 3d'{}^1F$	1.0	6
6800	$3p'{}^3P - 3d'{}^3D$	1.0	9

Oxygen VI—Notation and analysis are from Edlén (1933—1 and 1933—2). The calculations were carried out exactly as for Carbon IV. All the values of $c\sigma^2$, except for the transition $3s - 3p$, are extrapolated values. The adopted temperature was 100,000° K.

TABLE 16

SPECTRUM OF OXYGEN VI

λ	Combination	$\log c\sigma^2$	I
3314	$6P-7D$	0.89	100
3426	$6D-7F$	0.90	89
3433	$6F-7G$	0.54	39
3434	$6G-7H$	0.45	31
3434	$6H-7I$	0.11	14
3438	$6F-7D$	0.28	21
3509	$6D-7P$	0.62	43
3622	$6P-7S$	0.92	75
3819	$3S-3P$	1.034	11,600
4751	$7S-8P$	1.72	94
5112	$7P-8D$	1.08	16
5279	$7D-8F$	1.20	18
5289	$7F-8G$	0.84	8
5291	$7G-8H$	0.24	2
5291	$7H-8I$	0.31	2
5291	$7I-8K$	0.19	2
5298	$7F-8D$	0.70	6
5410	$7D-8P$	0.99	10
5602	$7P-8S$	1.17	14

The lines corresponding to the transition $3S - 3P$ have an excitation potential forty-eight volts lower than the other lines; at a temperature of 500,000° K they should still be seven times as strong as the next strongest lines.

Sodium I—The only lines of Na I that need to be considered are the D lines, which are so close together in wave-length that they were treated as one line with an arbitrary intensity in the final synthesis.

TABLE 17

ADJUSTMENT OF INTENSITIES WITHIN MULTIPLETS OF IRON II

λ	Combination	I^2_{lab}	I_{calc}
4129	$b^4P_{2\frac{1}{2}} - z^4D^\circ_{1\frac{1}{2}}$	9	6
4173	$2\frac{1}{2}$ $2\frac{1}{2}$	64	56
4233	$2\frac{1}{2}$ $3\frac{1}{2}$	(121)	250
4273	$1\frac{1}{2}$ $\frac{1}{2}$	9	11
4303	$1\frac{1}{2}$ $1\frac{1}{2}$	64	67
4352	$1\frac{1}{2}$ $2\frac{1}{2}$	(81)	131
4385	$\frac{1}{2}$ $\frac{1}{2}$	49	52
4417	$\frac{1}{2}$ $1\frac{1}{2}$	49	52
4473	$b^4F_{2\frac{1}{2}} - z^4F^\circ_{1\frac{1}{2}}$	4	8
4489	$3\frac{1}{2}$ $2\frac{1}{2}$	16	10
4491	$1\frac{1}{2}$ $1\frac{1}{2}$	25	31
4515	$2\frac{1}{2}$ $2\frac{1}{2}$	49	40
4520	$4\frac{1}{2}$ $3\frac{1}{2}$	49	8
4534	$1\frac{1}{2}$ $2\frac{1}{2}$	4	8
4556	$3\frac{1}{2}$ $3\frac{1}{2}$	64	59
4583	$2\frac{1}{2}$ $3\frac{1}{2}$	9	10
4629	$4\frac{1}{2}$ $4\frac{1}{2}$	49	89
4666	$3\frac{1}{2}$ $4\frac{1}{2}$	4	8
4508	$b^4F_{1\frac{1}{2}} - z^4D^\circ_{\frac{1}{2}}$	64	45
4523	$2\frac{1}{2}$ $1\frac{1}{2}$	(81)	72
4542	$1\frac{1}{2}$ $1\frac{1}{2}$	16	18
4549	$1\frac{1}{2}$ $2\frac{1}{2}$	(100)	110
4576	$2\frac{1}{2}$ $2\frac{1}{2}$	16	24
4584	$4\frac{1}{2}$ $3\frac{1}{2}$	(121)	161
4620	$3\frac{1}{2}$ $3\frac{1}{2}$	9	18
5198	$a^4G_{2\frac{1}{2}} - z^4F^\circ_{1\frac{1}{2}}$	36	30
5234	$3\frac{1}{2}$ $2\frac{1}{2}$	49	40
5276	$4\frac{1}{2}$ $3\frac{1}{2}$	49	54
5317	$3\frac{1}{2}$ $2\frac{1}{2}$	64	70
5326	$3\frac{1}{2}$ $3\frac{1}{2}$	4	6
5425	$4\frac{1}{2}$ $4\frac{1}{2}$	4	5
4924	$a^6S_{2\frac{1}{2}} - z^6P^\circ_{1\frac{1}{2}}$	(144)	96
5018	$2\frac{1}{2}$ $2\frac{1}{2}$	(144)	144
5169	$2\frac{1}{2}$ $3\frac{1}{2}$	(144)	192
6148	$b^4D_{1\frac{1}{2}} - z^4P^\circ_{\frac{1}{2}}$	4	2
6149	$\frac{1}{2}$ $\frac{1}{2}$	4	2
6238	$1\frac{1}{2}$ $1\frac{1}{2}$	1	3
6248	$2\frac{1}{2}$ $1\frac{1}{2}$	9	6
6417	$2\frac{1}{2}$ $2\frac{1}{2}$	1	3
6456	$3\frac{1}{2}$ $2\frac{1}{2}$	9	12

Calcium II—The H and K lines are the only important lines of Ca II; the intensity of K was adopted as twice that of H.

Iron II—The spectrum of ionized iron is not susceptible of the type of treatment that has been used for the other spectra. Accordingly the adopted intensities were based on the laboratory intensities given by Dobbie (1938), which are on a scale from 0 to 20. For multiplets in the visual region, the tables given by Russell (1936) were used to calculate the theoretical intensities of individual lines. The intensities were assumed to be proportional to the strengths (the effects of frequency and excitation level being here neglected). For intensities up to and including 8 on Dobbie's scale, the square of the laboratory intensity was found to be very nearly proportional to the strength. However, for stronger lines, the intensities thus deduced from Dobbie's values fall below the theoretical strengths. The relevant data are given in Table 17, where the scale was established by means of the laboratory intensities for lines of intensity 8 and less, and the values for the stronger lines were deduced from the theory. For the shorter wave-lengths the corresponding calculation was not made, but probably no appreciable errors are introduced from this cause for wave-lengths greater than λ3300. Table 18 contains the adopted

TABLE 18

Grouped Intensities for Spectrum of Iron II

λ	I	λ	I	λ	I	λ	I
3003	46	3390	14	4051	3	5018	31
3021	1	3418	8	4074	4	5033	4
3039	13	3440	10	4124	5	5064	2
3060	39	3464	21	4176	26	5100	5
3077	57	3484	5	4233	51	5148	2
3113	41	3495	32	4266	4	5169	40
3133	25	3507	5	4301	21	5202	8
3145	6	3538	4	4314	4	5238	10
3154	31	3564	3	4322	2	5270	18
3167	59	3621	21	4355	33	5287	3
3183	100	3725	3	4385	11	5317	19
3194	64	3754	30	4417	11	5363	5
3212	55	3783	3	4452	6	5398	2
3229	53	3823	12	4490	10	5424	4
3243	54	3836	2	4517	34	5470	3
3259	64	3861	2	4549	43	5505	2
3278	34	3907	6	4583	40	5533	4
3294	27	3937	11	4629	27	5568	1
3303	7	3969	5	4664	2	6148	1
3325	16	4003	4	4724	6	6245	2
3363	8	4026	7	4924	20	6447	3

intensities obtained by summing the squares of the Dobbie intensities (corrected as above in the visual region) throughout small, arbitrarily chosen, intervals of wave-length.

Mechanical Procedure

Most of the available information on the spectra of supernovæ is contained in the paper by Minkowski (1939). He reproduces microphotometer tracings made from unwidened spectra, which have been taken with several different spectrographs and on various emulsions. Because of the nature of the microphotograms, it is not possible for us to reduce them even to a relative scale of intensities. Because, moreover, the effects of prismatic dispersion and plate sensitivity distort the original intensity distribution enormously, no direct comparison between theory and observation is possible. The most satisfactory alternative was for us to reduce the theoretical intensity distributions to approximately the system of the tracings, by including as well as possible the effects of dispersion and sensitivity, but neglecting the very complicated effects of the characteristic curves, exposure times, and spectrum widths. The combination of spectrograph and emulsion most frequently used by Minkowski was spectrograph "*f*" (called "*e*" when used with the 100-inch reflector) with Agfa Supersensitive Panchromatic film. He reproduces in his paper microphotometer tracings of four spectra of ζ Aquilæ, spectral type B9n, with known relative intensities. From these tracings a mean characteristic curve was deduced. On the assumption that the star radiates like a black body with a temperature of 12,000° K, we then calculated values of the product of relative sensitivity, transmission, and dispersion (angstroms per millimeter) over the wave-length range from λ3800 to λ6800. Numerical values of this product divided by wave-length are given in Table 19. Entries corresponding to wave-lengths between λ3800 and λ3000 were made by extrapolating the observed values by a smooth curve that reached zero at λ3000.

In order to carry out our theoretical synthesis from the relative intensities derived in the previous section it is necessary to assume a form for the emission line profiles. The observational material for the spectra of supernovæ shows only that the bright lines are very broad; a rough upper limit for the breadth can be inferred from the steepness of the sides of some of the features in the yellow and orange regions of the spectrum. Little indication of the

TABLE 19

Adopted Correction for Sensitivity and Dispersion

λ	Factor	λ	Factor	λ	Factor
3000	0.00	4300	1.02	5600	0.86
3100	0.02	4400	1.05	5700	1.00
3200	0.04	4500	1.05	5800	1.13
3300	0.08	4600	1.00	5900	1.26
3400	0.12	4700	0.96	6000	1.39
3500	0.16	4800	0.88	6100	1.64
3600	0.21	4900	0.62	6200	2.13
3700	0.29	5000	0.45	6300	2.50
3800	0.40	5100	0.41	6400	2.74
3900	0.55	5200	0.44	6500	2.69
4000	0.71	5300	0.51	6600	2.01
4100	0.85	5400	0.61	6700	0.34
4200	0.95	5500	0.73	6800	0.00

degree of symmetry can be obtained from the observations, but steep sides and rounded tops are suggested. The most convenient form that is consistent with the observations is a parabola with the major axis vertical and the vertex upward. A parabolic profile is not inconsistent with the idea that an expansion is one of the causes of the brightening of a supernova, the breadth of the lines being caused by a Doppler effect. Although the elementary theory

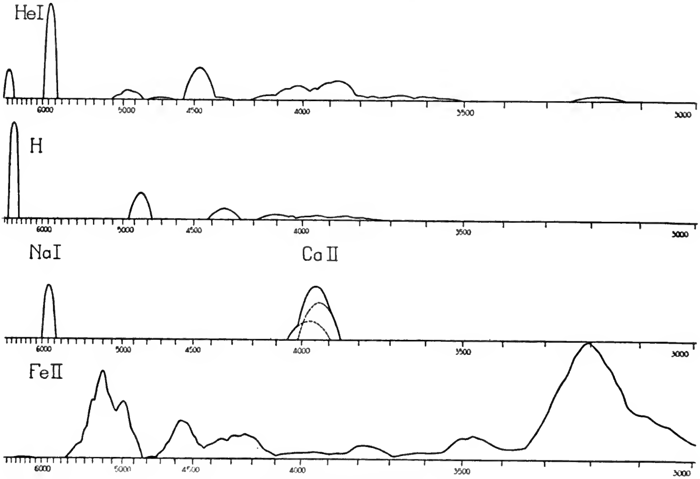


Fig. 1a. Fundamental spectra.

of ejection, as formulated for a Wolf-Rayet star by Beals (1930), leads to a flat-topped profile, the more refined discussion by Chandrasekhar (1934), which takes account of acceleration, allows in one case for a profile that is nearly symmetrical and not very far from parabolic. For ordinary novæ, the variety of profiles represented observationally is tremendous, and it would be impossible to select a single one as typical, although a parabola would be a fair average. In conformity with the evidence assembled by Whipple (1939) we have adopted a velocity of expansion of 6000 km/sec, so that the bases of the parabolas correspond to a range in velocity of 12,000 km/sec. The same velocity has been used for all the predictions, although there are some indications that the velocities may decrease in the late stages.

The individual integrated spectra are shown in Figs. 1a to 1d.

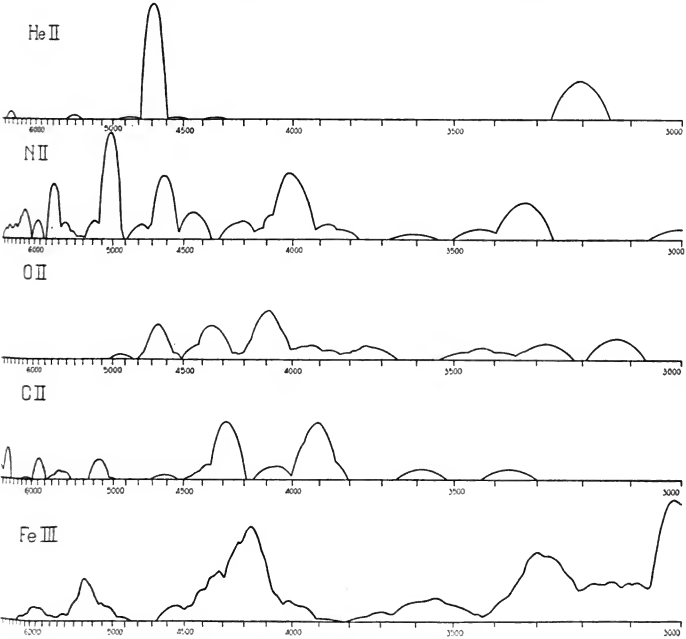


Fig. 1b. Fundamental spectra.

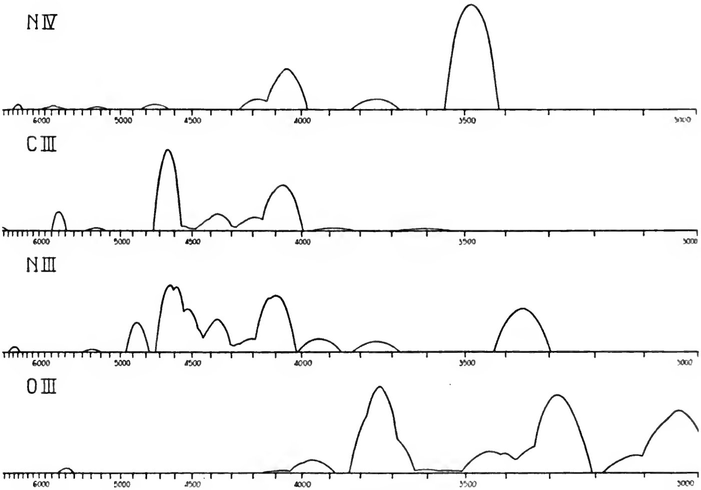


Fig. 1c. Fundamental spectra.

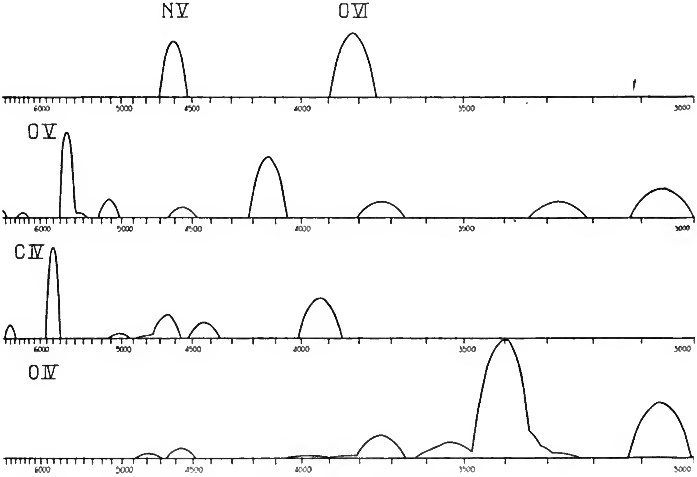


Fig. 1d. Fundamental spectra.

The intensities taken from tables 2 to 18 are used as the heights of the vertices of the parabolas in the summations. The bases of the parabolas are proportional to the wave-lengths of the centers and are reduced to the prismatic dispersion scale. This method of drawing required the introduction of the wave-length as a reciprocal factor in the calculation of the correction for sensitivity and dispersion (Table 19) to be applied later. The effects of the corrections for sensitivity and dispersion can be seen from the 10,000° continuum reproduced in Fig. 3.

Integrations of the individual atomic spectra as well as their final integration into composite spectra were carried out by simple summation at intervals of wave-length small enough to provide the necessary resolving power. The atomic spectra were integrated graphically, and the corresponding ordinates were read off and recorded on strips of paper. Subsequent integrations and calculations were made from these strips numerically. An example of the graphical integration may be seen in Fig. 1a for the two lines of ionized calcium.

THE SYNTHESIZED SPECTRA

The atomic spectra were assembled into five groups in order of increasing ionization and excitation. The composition of these basic spectra, which we have designated A, B, C, D, and E, is shown in the first six columns of Table 20. The tabular entries

TABLE 20

ATOMIC COMPOSITION OF THE SYNTHETIC SPECTRA

Percentage of Energy from λ3800 to λ6800

Element	A	B	C	D	E	II	IIa	IIb	IIIa	IIIb	IIIc	IVa	IVb
H I	4.9	6.9				5.9	4.7	3.9	3.6	2.4	1.6		
He I	7.4	15.6				11.5	9.2	7.7	7.8	5.3	3.6		
He II		9.6	16.9			4.8	7.2	9.5	12.0	13.6	15.2	8.4	5.6
Na I	2.2					1.1	0.9	0.6	0.2	0.1			
Ca II	10.7					5.4	4.3	3.1	1.0	0.7			
C II		18.2				9.1	7.3	6.5	8.3	5.7	4.2		
C III			23.2				4.6	8.3	10.5	14.5	17.8	11.6	7.7
C IV				68.0								17.0	22.7
N II		24.3				12.1	9.7	8.7	11.0	7.6	5.6		
N III			57.2				11.4	20.3	26.1	35.8	44.1	28.6	19.0
N IV				24.0								6.0	8.0
N V					18.0							4.5	6.0
O II		8.6				4.3	3.4	3.1	3.9	2.7	2.0		
O III			2.7				0.6	1.0	1.2	1.7	2.1	1.4	0.9
O IV				8.0								2.0	2.7
O V					62.0							15.5	20.7
O VI					20.0							5.0	6.7
Fe II	74.8					37.4	30.0	21.3	6.8	4.7			
Fe III		16.8				8.4	6.7	6.0	7.6	5.2	3.8		
Sum	100	100	100	100	100	100	100	100	100	100	100	100	100

show the percentage of energy contributed to each of the basic spectra by the atomic spectra shown in the first column. The energy is summed for the wave-length interval λ3800 to λ6800, and is not corrected for the sensitivity factor of the emulsion; the tabular entries are therefore bolometric energies. The percentages were chosen on the basis of the observed spectra of novæ and supernovæ, and of astrophysical abundance. These composite atomic spectra, A, B, C, D, and E, uncorrected for dispersion and sensitivity, are shown in Fig. 2.

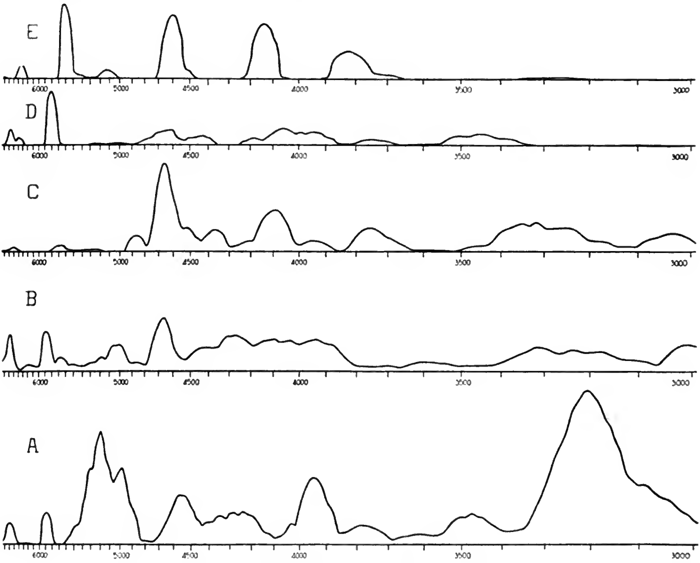


Fig. 2. Composite atomic spectra.

With each one of the spectra A, B, and C we associated continua of color temperatures 10,000° K, 16,000° K, and 28,000° K, respectively. In Fig. 3 is shown the continuum for 10,000° K by itself, and also combined with various proportions of spectrum A, all corrected for sensitivity and dispersion by the factors of Table 19. The corresponding bolometric energies contributed by

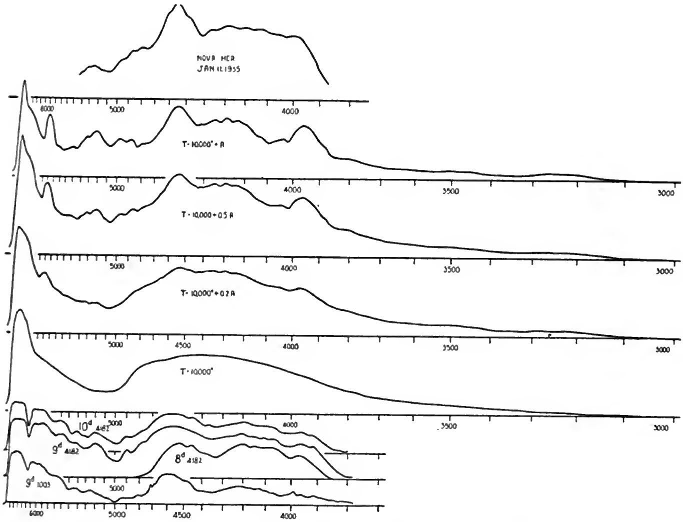


FIG. 3. Observed and synthesized supernova spectra—early stages.

the lines and the continuum are shown in the first four columns of Table 21.

TABLE 21

COMPONENTS OF THE SYNTHETIC SPECTRA

Percentage of Energy from λ3800 to λ6800

Components	0.2 A + 10,000°	0.5 A + 10,000°	1.0 A + 10,000°	II	IIa	IIb	IIIa	IIIb	IIIc	IVa	IVb
A	10.2	22.1	36.1	28.0	33.9	30.0	10.5	8.2			
B				10.0	12.0	13.3	18.8	14.6	13.6		
C					3.4	7.5	10.6	16.4	25.6	25.3	19.4
D										14.9	22.7
E										16.5	24.9
10,000°	89.8	77.9	63.9	49.6	29.9	19.8	18.6	14.5			
16,000°				12.4	15.0	16.5	23.3	18.1	16.9		
28,000°					5.8	12.9	18.2	28.2	43.9	43.3	33.0
Sum	100	100	100	100	100	100	100	100	100	100	100
Sum lines	10.2	22.1	36.1	38.0	49.3	50.8	39.9	39.2	39.2	56.7	67.0
Sum cont.	89.8	77.9	63.9	62.0	50.7	49.2	60.1	60.8	60.8	43.3	33.0

The synthesized spectra II, IIa, and IIb of Fig. 4, and IIIa, IIIb, and IIIc of Fig. 5 are compounded from the spectra A, B, and C, and their corresponding continua, and represent a series of spectra with increasing excitation, both of bright lines and of continuous background. The detailed atomic compositions are shown in Table 20, columns 7 to 12, inclusive, and the proportions of A, B, and C, and of bright lines to continuum are shown in Table 21, columns 5 to 10, inclusive.

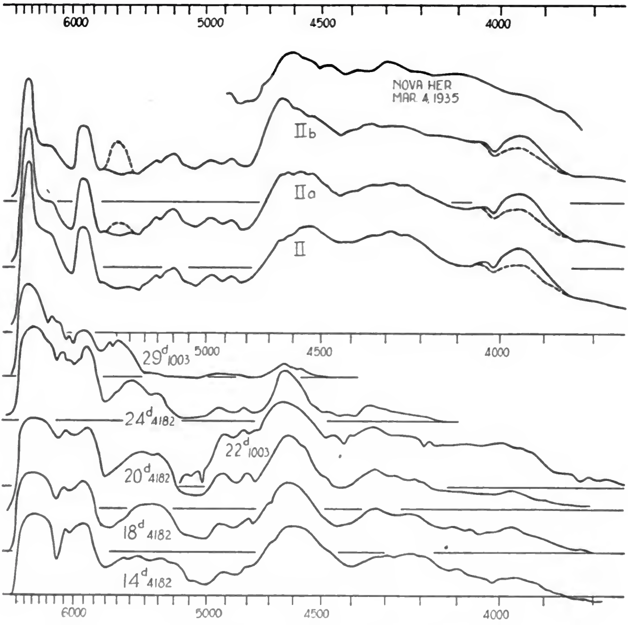


Fig. 4. Observed and synthesized supernova spectra—intermediate stages.

Spectra IVa and IVb of Fig. 5 represent still higher stages of excitation, being composed of spectra C, D, and E, and the continuum corresponding to C. Their compositions are shown in the last two columns of Tables 20 and 21.

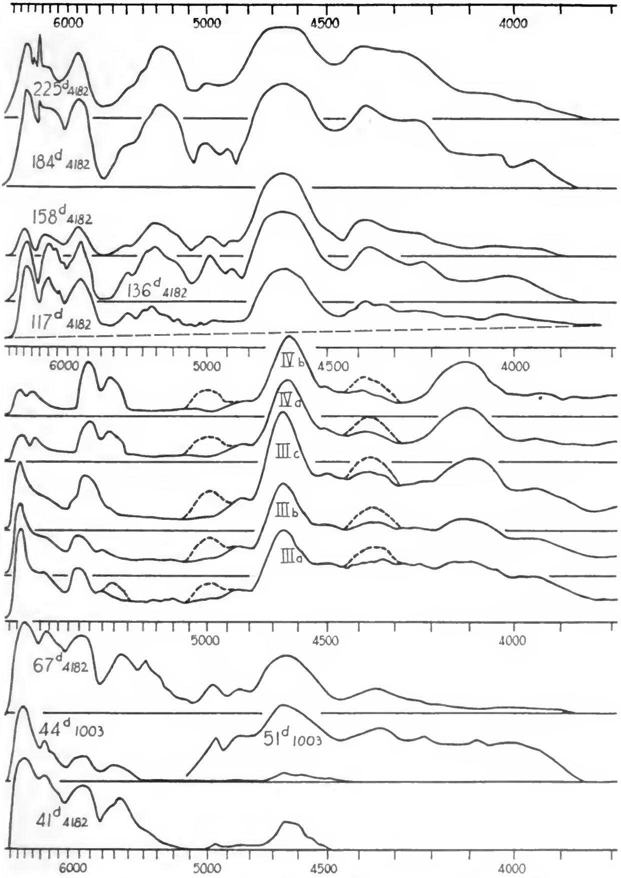


Fig. 5. Observed and synthesized supernova spectra—later stages.

COMPARISON WITH OBSERVATION

For purposes of comparison with the synthetic spectra, some of Minkowski's microphotograms are reproduced in Figs. 3, 4, and 5. The identification number of the supernova and the interval in days from maximum light are given for each curve in the figures. The diagrams in Minkowski's article were carefully traced on transparent squared paper, the curve being drawn with a view to eliminating plate grain, but preserving the essential details. The coordinates of the curve were then read off, and the curve replotted on the standard scale adopted for the synthetic spectra, which was exactly twice the average scale for Minkowski's reproductions for spectrograph "*f*." Most of the synthetic spectra are drawn with a larger vertical magnification than the microphotograms, because it was impossible to make an allowance for the effect of the characteristic curve on the microphotograms.

Intervals of 9 and 10 days after maximum light for the supernovæ are associated in Fig. 3 with spectrum A, of lowest excitation. The synthesized spectra are extended into the ultraviolet, to show the featureless character to be expected in this region. The three predicted spectra in Fig. 3 show the effect of varying the proportion of energy contributed by the bright lines and the continuum, and illustrate the workmanship of making the syntheses.

The spectra labelled "Nova Herculis," in an early stage at the top of Fig. 3 and in a later stage at the top of Fig. 4, were reproduced from the measures published by the authors (1937). The measured intensities of the bright lines were used; the lines were widened and replotted as for the synthesized spectra and superimposed on the original continuum, the hydrogen lines being reduced greatly in strength. The modified spectra of Nova Herculis thus demonstrate the marked changes produced in well known spectra by a broadening of the emission lines and a reduction of the strength of the hydrogen lines. The state of excitation of the emission lines in Nova Herculis had increased greatly in the interval between the two spectra shown. The general similarity between these spectra and some of the supernova spectra supports the authors' premise that the physical conditions are somewhat similar in character, differing principally in velocity of ejection, and that the state of excitation in the luminous atmospheres of the supernovæ increases with the time.

We can scarcely expect that the spectra of novæ and supernovæ will be identical when we consider the probable differences in atomic abundance and the huge difference in order of magnitude of the phenomena. The spectrum of Nova Herculis was deliberately chosen for the comparison because it was, like that of Nova Pictoris, very rich in ionized iron lines; in the spectrum of Nova Aquilæ, on the other hand, lines of this element were inconspicuous.

In Fig. 4, the synthesized spectra in stage II are associated with supernova spectra from 14 to 29 days after maximum. In this and subsequent spectra the predicted spectra in the ultraviolet are not reproduced, because of lack of observational material for the supernovæ. The relative intensities of the lines in the spectra adopted for the synthesis are given down to λ3000 in Tables 2 to 18. By far the strongest lines in the ultraviolet region are those of ionized iron, and the contribution made by these lines to the predicted spectra is shown in Figs. 1a and 3. In spectra II, IIa, and IIb the broken line represents the effect of reducing the intensity of ionized calcium to half the value given in Table 20, and also of including the forbidden line λ5577 of neutral oxygen.

Figure 5 compares the synthesized spectra in stages III and IV with supernova spectra from 41 to 67, and from 117 to 225 days after maximum. The broken curves in the predicted spectra show the effect of adding the forbidden lines λ4363 of O III, λ4958 and λ5007 of O III, and λ5577 of O I. The predicted spectra of highest excitation (IIIc, IVa, and IVb) do not agree satisfactorily with the observed spectra, and show the limitations on our method of prediction imposed by insufficient knowledge of high-excitation spectra, by the wide range of possibilities, and also by the unknown effects of the dilution of the primary radiation.

General Discussion

The red shift in the blue regions of the spectrum can be seen in Figs. 3, 4, and 5 for both the observed and predicted spectra. It appears to be a chance phenomenon arising from the effect of a progressively increasing excitation upon the bright-line spectra involved. An additional source of a small red shift is illustrated in Fig. 6 for the broadened lines $H\beta$ and $N_1 + N_2$, separately. For most photographic emulsions there is a rapid decrease in sensitivity with increasing wave-length in the blue-green regions. The effect of increasing line intensity with respect to the continuum, or of decreasing continuum with respect to the line intensity, will be to produce a spurious red shift for lines in this region of the spectrum. The second alternative is illustrated in the figure. Where the sensitivity gradient with wave-length or

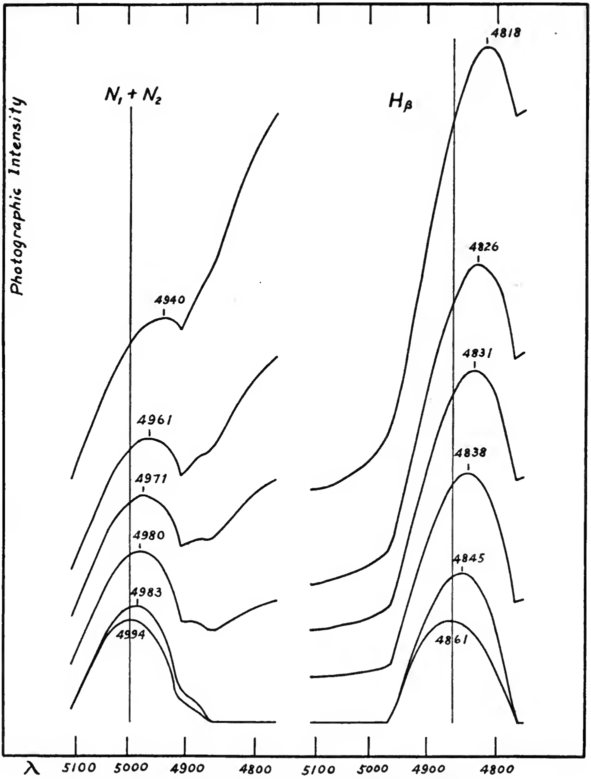


FIG. 6. Spurious red shift produced by changes in relative intensity of background and lines.

the relative change in intensity are reversed, a blue shift will result. No marked example of the conditions necessary to produce an appreciable blue shift is presented by the supernova spectrum.

In a recent note the authors (1940) have discussed the general problem of identifying the observed features in the spectra of supernovæ. It is unnecessary to repeat the detailed arguments. The present paper contains the fundamental material upon which the discussion is based. The conclusions, in addition to those already drawn above, may now be briefly summarized.

Near maximum light the spectrum of a supernova consists chiefly of a low-temperature (10,000° K ±) continuum, upon which are superimposed widened emission lines of common elements in states of relatively low excitation. Probably molecular absorption bands are also present in the early stages, cyanogen absorption near λ3800 and possibly titanium oxide absorption near λ6140.

With the progress of time, the state of excitation increases, together with the temperature of the continuum, while the intensity of the continuum diminishes, in the visual and photographic regions.

The narrow metastable lines of neutral oxygen at λ6300 and λ6363 that are observed in the later stages are ascribed to an interstellar source in the neighborhood of the supernova, excited by the pulse of high-intensity energy emitted near maximum light. Their low excitation at such a late stage, and their narrowness, are thus explained.

The evidence concerning the occurrence of other forbidden lines is not very conclusive, chiefly because it is difficult to predict their intensities, relative both to each other and to the permitted lines. This difficulty is particularly marked for Fe II and Fe III. Probably the inclusion of the [O III] lines improves the agreement slightly. There is no evidence for the occurrence of [N II] or of [O II].

The peculiar relative abundance of various atoms in the atmospheres of supernovæ, notably the paucity of hydrogen and quantity of iron, that is required by our synthesis presents a problem of great interest.

The authors are especially indebted to Dr. Leo Goldberg for advice with regard to the theory of line intensities, and to Drs. B. Edlén and P. Swings for use of their unpublished data on Fe III.

REFERENCES

BEALS, C. S. 1930. *M. N. R. A. S.*, **90**, 202.
CANNON, A. J. 1913. *Har. Ann.*, **76**, p. 37 and Plate 2.
CHANDRASEKHAR, S. 1934. *M. N. R. A. S.*, **94**, 522.
CONDON, E. U., AND SHORTLEY, G. H. 1935. *The Theory of Atomic Spectra*, p. 133.

Dobbie, J. C. 1938. *Ann. Sol. Phys. Obs. Camb.*, **5**, Part I, 1–59.

Edlén, B. 1933—1. *Nova Acta Reg. Soc. Sci. Upsala*, Ser. IV, Vol. 9, No. 6.

Edlén, B. 1933—2. *Zs. f. Ap.*, **7**, 378–390.

Edlén, B. 1935. *Zs. f. Phys.*, **93**, 726–730.

Fowler, A. 1926. *Proc. Roy. Soc.*, A, **110**, 476–501.

Freeman, L. J. 1928. *Proc. Roy. Soc.*, A, **121**, 318–343; (1929), *Proc. Roy. Soc.*, A, **124**, 654.

Goldberg, L. 1935. *Ap. J.*, **82**, 1–25; (1939), *Ap. J.*, **90**, 414–428.

Humason, M. L. 1936. *P. A. S. P.*, **48**, 110–113.

Humason, M. L., and Minkowski, R. 1940. *P. A. S. P.*, **52**, 146.

Johnson, W. A. 1936. *Har. Bul. 902*, 11–13.

Kuhn, W., Thomas, L. H. and Reiche, F. See Unsöld, A., *Physik der Sternatmosphären*, 1938, p. 190.

Menzel, D. H., and Pekeris, C. L. 1935. *M. N. R. A. S.*, **96**, 77–111.

Minkowski, R. 1939. *Ap. J.*, **89**, 143–204.

Moore, C. E. 1933. *A Multiplet Table of Astrophysical Interest*, Princeton, N. J.

Payne-Gaposchkin, C. 1936. *Ap. J.*, **83**, 245–251.

Payne-Gaposchkin, C., and Whipple, F. L. 1940. *P. N. A. S.*, **26**, 264–272.

Pickering, E. C. 1897. *Har. Circ.* No. 4.

Popper, D. M. 1937. *P. A. S. P.*, **49**, 283–289.

Ritchey, G. W., and Pease, F. 1917. *P. A. S. P.*, **29**, 211.

Russell, H. N. 1925. *P. N. A. S.*, **11**, 322; (1936), *Ap. J.*, **83**, 129–139.

Strohmeier, W. 1937. *Zs. f. Ap.*, **14**, 227; (1938), *Die Sterne*, **18**, 71.

Vorontsov-Velyaminov, B. 1931. *A. N.*, **242**, 208.

Whipple, F. L. 1939. *These Proceedings*, **81**, 253.

Whipple, F. L., and Payne-Gaposchkin, C. 1937. *Har. Circ.* No. 414.

PRODUCTION OF SECONDARY ELECTRONS BY ELECTRONS OF ENERGY BETWEEN 0.7 AND 2.6 MEV

GEORGE HORNBECK AND IRL HOWELL

Department of Physics, University of North Carolina

Abstract

The production of secondary electrons in the gas of a cloud chamber, by electrons with energy from 0.67 to 2.6 MEV, is studied, using 190 meters of track. The energies of the secondaries are obtained by measuring their ranges. Thus it is possible to obtain values of the cross-section for production of a secondary with kinetic energy T, greater than some chosen value T_0, but less than $T_p/2$, where T_p is the primary kinetic energy. These may be compared with values calculated from Möller's theory. If R is the ratio of the observed and calculated cross-sections (averaged over the various values of primary energy), we have:

T_0(EKV):	20	30	40
R :	1.18 ± .14	1.09 ± .16	0.84 ± .15

Thus the results are substantially in agreement with theory, just as those of Champion are in the lower T_p-range from 0.4 to 0.9 MEV. Williams and Terroux reported large discrepancies, obtaining R-values which varied with both T_0 and T_p, ranging, in fact, from about 1.6 to 2.9. We show that their results can probably be explained on the basis of two factors: (1) inclusion of secondaries with ranges so small that it is difficult to obtain their energies; (2) the number of secondaries of high energy is small, which leads to large statistical fluctuations.

We conclude that theory and experiment are essentially in agreement over the T_p-range from 0.4 to 2.6 MEV; the statement applies for values of T/T_p lying in the limited range from 0.02 to 0.1.

I. PREVIOUS WORK

At present it is important to obtain experimental data on the cross-section for the scattering of high-energy electrons by other electrons. There are many studies, both theoretical and experimental, on energy loss, range, straggling and ionization, for primary energies up to about 10 MEV.[1] These phenomena give indirect information about the electron-electron cross-section for small values of the relative energy transfer. On the other hand, direct experiments on the energy distribution of secondary electrons and the angular distribution of scattered primaries are the only ones which yield significant information about the cross-section for "close" collisions, *i.e.*, those in which a moderate

---

[1] See, for example, the energy loss measurements of Crane and his colleagues, especially Crane and Slawsky, *Phys. Rev.*, **56**, 1203 (1939); and of Laslett and Hurst, *Phys. Rev.*, **52**, 1035 (1937). Also the ionization measurements of Corson and Brode as reported in *Rev. Modern Physics*, **11**, 222 (1939). An excellent analysis of the situation as it existed in 1932 is given by Williams, *Proc. Roy. Soc.*, **135**, 108 (1932).

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 1, APRIL, 1941

fraction or a large fraction of the energy is transferred. In this paper we shall be concerned exclusively with experiments of this type. Let us consider first the high-energy region.

Anderson and Neddermeyer [2] have studied the energy distribution of fast negative secondaries, ejected from a plate in a cloud chamber by cosmic-ray particles. In one case, most of the primaries had energy $\gg$200 MEV, and secondaries up to 80 MEV were found; in 3 other cases, the primaries had energy $\gg$500 MEV and secondaries up to about 150 MEV were observed. In these experiments there is uncertainty as to whether an individual primary is an electron or a meson. This hinders comparison with theory, because it is necessary to make a correction for secondaries which fail to emerge from the plate; the influence of collision losses is fairly clear but the question is, what allowance should be made for loss of secondaries caused by radiative processes? Two extreme cases were therefore considered—no radiative loss, and full radiative loss for all particles, considered as electrons. The results show that it is probably fair to say that, regardless of the nature of the primaries, the observed and calculated cross-sections for production of high energy secondaries agree within a factor 2 over the whole range of the experiments. More detailed statements can hardly be made because of the uncertainties mentioned, and because the number of secondaries available is small (to wit, 26, 11, 13, and 36 in the 4 cases); as we shall see, the smallness of the cross-section constitutes the chief barrier to progress in all studies of this kind. Nevertheless, the above result is reassuring since it shows that the Möller formula (discussed in Section III) cannot be far wrong for either electrons or mesons at very high values of primary energy, and for average values of the ratio (secondary energy)/(primary energy) lying in the neighborhood of 0.04.

We turn now to the domain of much lower primary energies, only to find some puzzling discrepancies which have led us to make the experiments described in this paper. So far as we can find, there are only six experimental papers which give results of direct scattering measurements in the energy region from about 10^4 to 10^6 MEV, and the last of these was published in 1932. The six investigations, in chronological order, are due to Bothe,[3]

---

[2] Anderson, C. D., and Neddermeyer, S., "International Conference on Physics," Vol. 1, p. 171, 1934, Cambridge Univ. Press; *Rev. Modern Physics*, 11, 191 (1939).

[3] Bothe, W., *Z. Physik*, 12, 117 (1922).

Wilson,[4] Henderson,[5] Williams and Terroux,[6] Williams[7] and Champion.[8] With the exception of Henderson's work, all the data were obtained by studying side branches produced by fast electrons in the gas of a cloud chamber. In the domain of energy in which radiative loss can be neglected, the energy loss of an electron traversing a gas is principally due to excitation of the atoms and to the production of slow secondaries, the kinetic energy of which is of the order of 15 volts. Occasionally, however, the energy transfer is much larger, and when it is of the order of 10^4 volts, the range of the secondary in air at N. T. P. is several millimeters. Then it is possible to determine the energy transfer by using the range-energy relation. With suitable choice of the energies involved, it is also possible to study the angular distribution of the scattered primary electrons. Fig. 1 is a stereoscopic pair showing two events in which the side branches are suitable for range measurement; they are indicated by arrows and may be examined with a hand mirror placed perpendicular to the paper, between the two views.

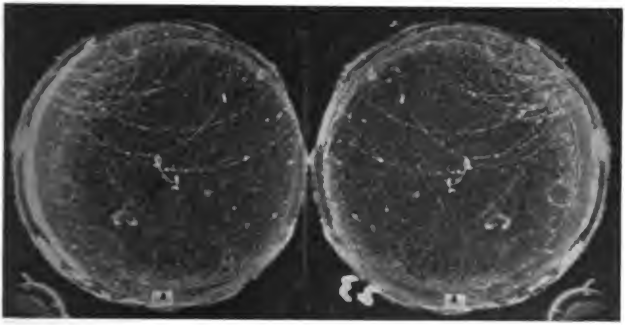


Fig. 1.

The frequency of such collisions is governed roughly by a formula contained implicitly in a paper by J. J. Thomson;[9] we shall call it Thomson's formula. According to this, the classical cross-section for an energy transfer between T_1 and T_2 is

---

[4] Wilson, C. T. R., *Proc. Roy. Soc.*, **104**, 192 (1923).

[5] Henderson, M. C., *Phil. Mag.*, **8**, 847 (1929).

[6] Williams, E. J., and Terroux, F. R., *Proc. Roy. Soc. A*, **126**, 289 (1929–30).

[7] Williams, E. J., *Proc. Roy. Soc. A*, **128**, 459 (1930).

[8] Champion, F. C., *Proc. Roy. Soc. A*, **137**, 688 (1932).

[9] Thomson, J. J., *Phil. Mag.*, **23**, 449 (1912); see also Bohr, *Phil. Mag.*, **25**, 10 (1913), and **30**, 581 (1913).

$$S_{\mathrm{cl}}(T_1, T_2) = \frac{2\pi e^4}{mv_p^2}\left(\frac{1}{T_1} - \frac{1}{T_2}\right), \tag{1}$$

where v_p is the primary velocity. This applies to the collisions of free electrons; corrections for the velocity of the atomic electron and for its binding energy are negligible in the range of primary energies considered here. One complication arises: one cannot tell whether a side branch of energy T is formed by the atomic electron, or by a primary which lost energy $T_p - T$. Therefore we agree that hereafter T shall mean the energy of the slower emergent particle, so that T is always less than $T_p/2$. Then the classical cross-section for production of a branch which has energy greater than a chosen value T_0, but less than $T_p/2$, is:

$$S_{\mathrm{cl}}(T_0, T_p - T_0) = \frac{2\pi e^4}{mv_p^2}\left(\frac{1}{T_0} - \frac{1}{T_p - T_0}\right). \tag{2}$$

This cross-section decreases rapidly as T_0 increases. Accordingly all experimenters using cloud chambers of ordinary size have been obliged to choose the ranges of primary and secondary energy in such a way that a reasonable number of events suitable for measurement could be accumulated. If T_p is of the order 10^4 to $3 \cdot 10^4$ EV, relatively large values of the fractional energy transfer, T/T_p, can be dealt with; but when T_p is 10^5 EV or more, events for which T/T_p is relatively large become so infrequent that their study in ordinary cloud chambers becomes prohibitively time-consuming. Considerations of this kind explain the paucity of quantitative data on this cross-section.

The pioneer cloud chamber studies of side branches were made by Bothe and Wilson without a magnetic field. The results showed that for the energies employed the frequency of side branches agrees with Thomson's formula in order of magnitude. Henderson's work was done with an ionization chamber. On account of the large contribution from nuclear scattering and the uncertainties due to scattering from solid parts of the apparatus the results throw no light on the present problem. Williams studied secondaries produced in oxygen by primaries having the low energy 20 EKV. They produce secondaries copiously, and by operating at pressures below atmospheric he was able to measure the frequency of branches with energies from 3,000 to 10,000 EV. Williams chose these conditions because he desired to test a scattering formula which Mott[10] derived on the basis of the Schrödinger equation, taking account of electron exchange. The difference between the Mott formula and the classical equation (2) is most pronounced in collisions for which T/T_p is large. Williams' results agree with Mott's formula within the limits of experimental error, about 10 per cent, while the most reliable portion of the data yields a cross-section 40 per cent below the classical value.

Williams and Terroux examined the side branches produced on 18 meters of track in oxygen at about two-thirds of atmospheric pressure, using primary electrons with energies from 0.13 to 1.6 MEV.[11] They examined all secondaries with energy over 7,500 EV. In their table of results, the observed secondaries are grouped in two ways, to facilitate study of the way in which the cross-section depends on T_0 and on T_p. First the events are classified into four groups, according to the value of T_0, without regard to the value of T_p. Second, all events for which T was greater than 10,000 EV were divided into groups, according to the value of T_p. In each case, the theoretical number of events was obtained from Thomson's formula, the only one available at that time. Comparison was made with theory by computing the ratio, "observed number over theoretical number," which we shall denote by R. The chief features of the results are these. (1) As T increases, R runs from 2 to 0.5, all values of T_p being considered. (2) As T_p increases from 0.13 to 1.6 MEV, R rises from 1.4 to 2.8, for the case of secondaries having energy greater than 10 EKV.

It is possible to compare the results with a relativistic quantum-theoretical formula based on Möller's results (see equation (3)), and we have done this. The effect is that all R-values are slightly increased, and we shall use these revised R-values in making comparisons with our own data (Section V). The above results are open to an obvious criticism. The number of events in each primary-energy group is small so that unfavorable fluctuations may be largely responsible for the variation of R with T_p. Therefore it seems best to lump the data for all values of T and T_p. We have done this, and find that the experimental average cross-section for *all* the events (72 in number) is 2.1 ± 0.25 times the value predicted by Möller's formula. Terroux's energy distribution curve was employed in making these computations, and the estimate of error is merely the standard deviation due to fluctuations in the occurrence of secondaries. We may say at once that these results are in sharp contradiction with the results of Champion, and with our own.

---

[10] Mott, N. F., *Proc. Roy. Soc. A*, **126**, 259 (1930).

[11] These primaries were described as beta particles from RaE, but since the upper limit of the RaE spectrum lies at about 1.3 MEV it is necessary to consider the origin of the particles of higher energy. Later (*Proc. Roy. Soc. A*, 131, 90 (1931)), Terroux published an account of this spectrum, based on Madgwick's data up to 0.66 MEV, and on his own data, above that value. His work was done with the apparatus used by Williams and Terroux; the field was 500 gauss, and accordingly "the probable error in the measurement of $H\rho$ for any single fast track is of the order of 9 per cent." Now we find that up to nearly 1.3 MEV the curve of Terroux agrees well with the mean of several recent and reliable distribution curves; the faster particles form a weak "tail." Thus it appears unlikely that his magnetic field measurements were seriously in error; it is far more probable that the tail arose from multiple scattering, from the 9 per cent probable error in $H\rho$ and from the presence of particles belonging to the cosmic rays and the radioactive background. Observations in our own laboratory show that such particles may be far from negligible in a six-inch chamber operating periodically.

Champion studied the deflection of beta particles in the energy range 0.38 to 0.9 MEV, scattered by electrons in nitrogen. In 650 meters of track, he obtained 250 branches for which the angle of scattering lay between 10° and 40°. The angular distribution is described by Table I, where the error-estimates again are based on fluctuations alone. There is some indication that the observed cross-section is below the calculated one and that the discrepancy increases with secondary energy; still the limits of fluctuation error are such that we cannot speak of a definite disaccord with theory.[12] In a rough way, the apparent deficit at higher angles corresponds to a deficit of high-energy secondaries, because, for a primary electron of definite energy, a given angle of scattering corresponds to the production of a secondary of given

TABLE I

ANGULAR DISTRIBUTION OF SCATTERED RaE BETA PARTICLES, AFTER CHAMPION

Deflection, Degrees	Observed No.	Calculated No.	R
10–20	214±15	230	0.93
20–30	26± 5	30	0.87
30–max	10± 3	13	0.77
10–max	250±16	273	0.916

---

[12] The formula employed by Champion involves $(e^2/mc^2)^2$. It is to be presumed that he calculated this in the form $e^2(e/m)^2/c^4$, using the value $e = 4.77\ 10^{-10}$ which was currently accepted in 1932; no definite statement is made, but if this assumption be correct, use of the correct e-value will reduce Champion's R's by 1.25%, thereby increasing the discrepancy with theory.

energy. In other words, there is a relation between the angular distribution $f(\theta)d\theta$ and the energy distribution $g(T)dT$. It would be possible to calculate the secondary energy distribution which corresponds to Champion's published data, but this would be very time-consuming and is scarcely worth while, for we see that the average result would necessarily be in close agreement with Möller's formula for $g(T)$.

To summarize the essential points, Williams and Terroux obtained twice the Möller cross-section for primaries in the range 0.13–1.6 MEV, but Champion's angular distribution agrees well with Möller's calculations over the range 0.4–0.9 MEV. It is inherently probable that Champion's result is more trustworthy than that of Williams and Terroux because of the smaller statistical error and because of the difficulty of measuring the ranges of short secondaries. More extensive cloud chamber observations are needed, both to clear up the discrepancy and to extend the results to higher primary energy.

II. NATURE OF THE PRESENT EXPERIMENTS

In this paper we present data covering the range of primary energies from 0.67 to 2.6 MEV, and the range of secondary energies from 10 EKV up to the highest value encountered, namely about 250 EKV. The desideratum in such work is a complete quantitative study of the cross-section over a wide range of values of T_p and of T_0/T_p, with special emphasis on high values of T_0/T_p. The present experiments show that the results of Williams and Terroux cannot be correct; they also test equation (3), Section III, for primary energies much higher than those employed by earlier workers.

The primaries were recoil electrons produced in the walls of a nitrogen-filled chamber, by gamma rays from a mesothorium source containing a small percentage of radium. Thus the effects of the hard gamma rays from ThC″ are prominent. We have chosen to measure the energy distribution of the branches for our viewing apparatus is better adapted to radius and range measurements than to deflection measurements. Furthermore, a large body of photographs obtained by Dr. C. C. Jones was available and these were well adapted to radius measurements because they were obtained in a field of 1,200 gauss. In view of the strong variation with T_p which Williams and Terroux found in the range they covered, it appeared desirable to study the dependence of [illegible] in T_p even though theory indicates very slight [illegible] high velocities.

[illegible] RELATIVISTIC FORMULA FOR THE ELECTRON-ELECTRON CROSS-SECTION

[illegible] the electron-electron collision at high energies [illegible] Bethe and by Wolfe,[13] but these authors neg[illegible] of potentials. The effect of retardation was [illegible] who states that his two-electron Hamiltonian [illegible] c^{-2} inclusive. Uncertainty about higher [illegible] when one writes the classical Hamiltonian, [illegible] depend on the "structure" assumed for the [illegible] calculated the electron-electron cross-section [illegible] known to be equivalent to the use of Breit's [illegible] c^{-2} inclusive. His final formulas giving [illegible] production of a branch with energy in the [illegible] are presumably the most accurate ones avail[illegible] applies when T is large compared with [illegible] the atomic electron and small compared with [illegible] for use in the present work.

[illegible] Möller's differential cross-section from T_0 up to [illegible] the following cross-section for production of a [illegible] superior to T_0 (and of course, inferior to $T_p/2$):

$$[illegible] \left(\frac{1 - 2A_0}{1 - A_0} \right)$$

$$[illegible] - g)A_0 \frac{1 - A_0}{1 - 2A_0} \log \left(\frac{1 - A_0}{A_0} \right) \tag{3}$$

$$+ gA_0\left(\tfrac{1}{2} - A_0\right) \left(\frac{1 - A_0}{1 - 2A_0} \right) \Bigg].$$

---

[13] [illegible] Physik, **5**, 325 (1930); Wolfe, H. C., *Phys. Rev.*, **37**, 591 [illegible]

[illegible] Rev., **34**, 553 (1929).

[illegible] der Physik, **14**, 531 (1932). Carlson and Oppenheimer, [illegible] have made similar calculations. Möller's formula (76) and [illegible] Oppenheimer give identical results when $T_p \gg mc^2$, but for [illegible] the results are far from identical. Specifically, when [illegible] from each other by amounts ranging from 13 per cent to [illegible] the value of T/T_p. For $T_p = 2.5$ MEV the discrepancies [illegible] per cent. Indeed, Carlson and Oppenheimer describe their [illegible] to high velocity impacts, but they do not say *how* high; [illegible] portion of the discrepancy is due to the absence of a factor [illegible] the Carlson-Oppenheimer formula; this factor would be [illegible] Thomson's formula in the classical range; Möller's formula [illegible]

Here T_0 is in mc^2 units; $\beta_p = v_p/c$; $A_0 = T_0/T_p$; $g = [(\gamma - 1)/\gamma]^2$, and $\gamma = (1 - \beta_p^2)^{\frac{1}{2}}$. Also $r_0 = e^2/mc^2$ and we have employed the values $2\pi r_0^2 = 4.983\ 10^{-25}\ (1 \pm .0007)\ \text{cm}^2$; 1 mc^2 unit $= 0.5113\ (1 \pm .0002)$ MEV, which are based on $e = 4.8029\ 10^{-10}$, $e/m_0 = 5.2690\ 10^{-17}$, and $c = 2.99776\ 10^{10}$.

The terms outside the square bracket in (3) are identical with those in Thomson's equation (2), so the second and third terms in the bracket represent the effect of passage from non-relativistic Newtonian mechanics to the best available quantum-mechanical formulation. For the high primary energies and relatively low secondary energies employed in our work, these two terms are small compared with unity. They seldom contribute more than 10 per cent to the cross-section.

Working Formula.—Since we have to work with a range of primary energies, we must average formula (3) between fixed limits of T_p, say τ and $\tau + \Delta\tau$. We assume the energy spectrum is given as a hodograph of flat topped blocks (Fig. 2), so that in averaging over one of these, we give all energy values equal weight. The result contains integrals which have not been tabulated, so a good approximation is desirable. For our own data, it is satisfactory to average only over the term β_p^2 outside the bracket and to replace A_0 and g by their average values. The result is:

$$\begin{aligned} \bar{S}(T_0, T_p) &\cong \frac{2\pi r_0^2}{T_0}\left(\frac{1 - 2\bar{A}_0}{1 - \bar{A}_0}\right)\bar{B}(1 + f); \\ f &\equiv \frac{1}{2\Delta\tau}\log_e\left(\frac{1 + (\Delta\tau/\tau)}{1 + [\Delta\tau/(\tau + 2)]}\right). \end{aligned} \tag{4}$$

Here $\bar{B}$ stands for the result of putting $\bar{A}_0$, $\bar{g}$, in place of A_0 and g in the bracket, and all the energies are in mc² units.[16]

IV. TECHNIQUE

Our work was done with a stereoscopic viewing and measuring instrument recently described by Jones and Ruark.[17] Full details of the $H\rho$ determinations are given in their paper. The instrument is capable of measuring the components of momentum parallel and perpendicular to the field, and can deal accurately with tracks which make an angle θ of 30 or 40° with a plane perpendicular to the field, but the computations applying to large values of θ are time-consuming. Because of the large number of tracks to be measured, we have used only those primary tracks

---

[16] This formula was not quite good enough for the recomputation of the Williams-Terroux data mentioned in Sec. I; in that case we used graphical methods.

[17] Jones, C. C., and Ruark, A. E., *Proc. Amer. Phil. Soc.*, **82**, 253 (1940).

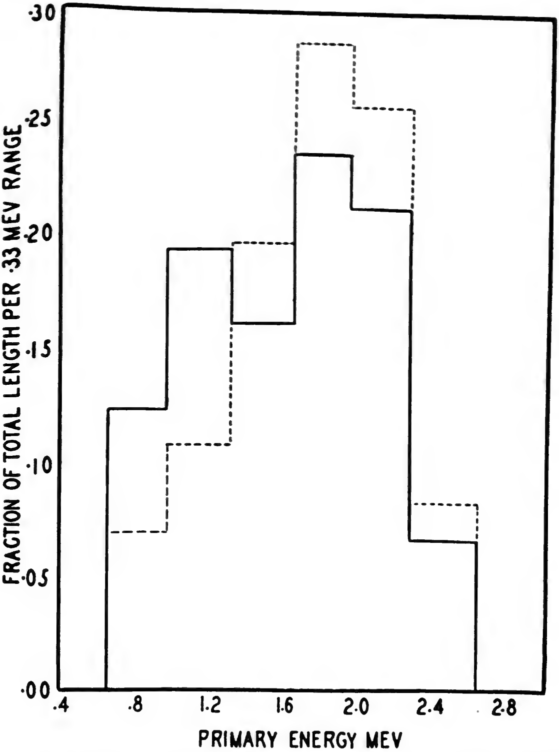


Fig. 2. Energy spectrum of the primary electrons. The solid and dotted curves refer to the data in Table II *A* and Table II *B*, respectively.

for which tan θ is less than 0.3. With this restriction, a small correction can be applied to the energy of a group of primary tracks to take account of both the momentum component parallel to the field and a geometrical error, caused by the fact that the helical tracks are projected on a flat ground-glass viewing plate. The appropriate corrections were determined by careful measurement of the actual configuration of 80 representative tracks.

For present purposes it is not necessary to discuss the effect of multiple scattering on either the energy distribution curve (Fig. 2) or the results (Fig. 3) for the dependence of the electron-electron cross-section on T_p. In a sufficiently large aggregate, the effects of multiple scattering on these quantities will be small for two reasons. First, the number of tracks thrown from one energy-class to the next higher one by scattering will be roughly compensated by the number which are thrown in the other direction. Second, many tracks have apparent energies which lie so far from a boundary between energy-classes that the chance that they really belong to another class is negligible.

The energies of the side-branch electrons were determined from their ranges. In measuring the length of a branch several centimeters long it was necessary to divide it up into reasonably straight portions, measuring each portion separately. With the mid-point of the segment to be measured in good focus on the ground-glass plate, the component normal to H was measured with a glass scale, while the component parallel to H was determined by a focusing method.[17] The chord was computed and, if necessary, the difference between arc and chord was estimated visually.

Since the cross-section depends critically on the branch energy, we shall have to consider the range-energy relation with care.

Nuttall and Williams [18] found that Whiddington's law is well verified in several gases for electrons of energy 5 to 27 EKV; that is, $T = Ar^{\frac{1}{2}}$, where T is the energy of the secondary in EKV, r is the range of the branch in cm, and $A = 22.3$ for nitrogen. They also show that the stopping power of a gas or gas mixture depends only on the electron density; this was verified within 1.4 per cent for several gases up to 27 EKV. From 27 to 50 EKV, deviations from Whiddington's law have been found, but they are slight. Von Droste [19] has given a universal range-energy curve, based on all available data for several gases up to 50 EKV; he converted all ranges into equivalent mean ranges in air at 15° C and 760 mm. It will be useful to give a few numbers taken from the curve:

---

[18] Nuttall, J. M., and Williams, E. J., *Phil. Mag.*, **2**, 1109 (1926).

[19] von Droste, G. Fr., *Z. Phys.*, **84**, 17 (1933).

T in EKV	r_{air} in cm
10	.22
20	.77
30	1.63
40	2.73
50	4.07

Now our photographs were obtained in nitrogen, with a liquid mixture of 60 per cent ethyl alcohol and 40 per cent water. The vapor pressures of the alcohol and water do not follow the laws of ideal mixtures, and corrections for the deviations are quite necessary. The actual vapor pressures of these two constituents at any temperature can be read from curves recently published by Gautier and Ruark.[20] The total pressure of the gas-vapor mixture is known from measurements taken at the beginning of each run. With the chamber in the expanded condition and in thermal equilibrium, it was usually about 78 cm. The nitrogen pressure is determined by difference, so we can compute the electron density of each constituent of the chamber gas, after expansion. We then assume (on the basis of Nuttall and Williams' results, as well as theory) that range is inversely proportional to electron density. Thus we get the range in air at 15° C and 760 mm, and take the energy from von Droste's curve. It appears better to use this curve than to rely on theoretical formulas, for the data of all observers except one agree with the curve within about one per cent, while the available formulas depart from the experimental results by amounts of the order 3 to 10 per cent.

Criteria for Selection of Tracks.—In taking the photographs used in this work, the source of gamma rays was enclosed in a thick lead housing with a narrow channel extending toward the chamber wall. The source was mounted on a strong electromagnet which was actuated about the time the expansion was complete, pushing the source into line with the channel. This shutter arrangement brought about a substantial increase in the percentage of post-expansion tracks. Nevertheless, in any series of electron photographs obtained with a gamma-ray source there will be a number of diffuse tracks which entered the chamber too early and a number of weak ones which entered too late, near the end of the sensitive period. Thus it is essential to adopt objective criteria as to the inclusion or exclusion of such tracks in advance of the measurements. The criteria depend on the purpose of the work and in any case they should be so chosen that decisions as to the inclusion of tracks are reduced to a minimum number. Also, tracks which lie in positions unfavorable for accurate measurement should be excluded.

---

[20] Gautier, T. N., and Ruark, A. E., *Phys. Rev.*, **57**, 1040 (1940).

Our criteria are these: (1) To minimize error in the total track length, the measurable length of each primary track should not fall below a definitely chosen value. For particles of energy less than 1 MEV, this value was 4 cm; for all others it was 5 cm. The primary track should be free from detectable small angle scattering over a consecutive length of at least 5 cm, *if* it is necessary to ascertain whether the primary lies within the energy range under consideration; otherwise, uniformity of radius is unimportant.

(2) The tangent of θ shall be less than or equal to 0.3. In the early part of the work the value 0.17 was used instead, but it was found that applications of this criterion were too frequent. Accordingly the figure was raised to 0.3. Because of the geometry of the chamber and the angular distribution of the recoil electrons employed, the number of applications dropped at once. This increased the speed of measurement and certainly resulted in greater objectivity of the data.

(3) Portions of a primary track obscured by other tracks should be omitted, though unobscured portions should be included. This trouble is infrequent.

(4) If a track primary becomes faint in the region of warm air very close to the chamber wall it should be considered to extend to the wall, for it actually does so, and it is probable that any slow secondary produced in this region can usually be detected with ease because of its great density of ionization. If this convention were not adopted, the cross-section would be overestimated.

(5) To minimize the percentage of pre-expansion tracks, and to increase the accuracy of curvature measurements, tracks the images of which are wider than the arbitrary value 0.8 mm are excluded; the width is determined by comparison with a standard ink line ruled on a transparent sheet.

(6) It must be probable from the density of ionization, the initial direction and the shape of a branch, that it is one which starts on the primary, rather than a slow electron which starts in the gas of the chamber and happens to end on the primary. This

criterion cannot easily be applied to branches only a few mm in length, but occasionally it has served to eliminate a false branch of considerable length. Often it is possible to distinguish a real branch from a false one by looking for the expected deflection of the primary. Experiments were made to determine the chance that a slow electron originating in the gas will terminate its path inside a cylinder of radius 2 mm surrounding a primary track; a very liberal estimate of this chance is 0.016. At least half of these accidentals can be eliminated by considering the angle they make with the primary and by looking for the deflection of the primary, so the error in cross-section due to accidentals can hardly be more than 0.8 per cent.

(7) The primary and the branch must be of the same age, as far as can be determined visually.

V. RESULTS

Our data are divided into two main series obtained with different purposes. Their essential characteristics may be described as follows:

Series	T_p Limits, MEV	T_0, EKV	Range of Branch, cm	Meters of Track
I	0.67–2.64	12, 20, 30, 40	0.3, 0.8, 1.8, 3.0	105.6
II	1.3 –2.64	20, 30, 40	0.8, 1.8, 3.0	84.6

In Series I, we employed a wide range of primary energies to provide good overlap with the values covered by previous workers. Nevertheless, the lower limit is so high that the average $H\rho$ value for a reasonable number of tracks is not seriously affected by multiple scattering. In this series the energies of all primaries and branches were determined individually so that the electron-electron collisions could be classified according to values of T_p and also according to values of T. In this series we measured all branches with ranges of 0.3 cm or more, because we wanted to duplicate the procedure of Williams and Terroux in this respect, and thereby get an indication as to whether the inclusion of such short tracks leads to large values of R, such as they observed.

In Series II, the lower limit of T_p was put at 1.33 MEV, and the lower limit of T_0 at 20 EKV, corresponding to a range of about 0.8 cm. This procedure avoids inaccuracies which may arise in the measurement of very short branches, increases the amount of data in the range of high T_p, not covered by others, and also builds up the statistics for higher values of T_0/T_p, where the cross-section is small. In this series, the dependence of the cross-section on T_p was not studied because the results of series I showed that the pronounced anomaly found by Williams and Terroux near $T_p = 1$ MEV has no foundation in fact (see Fig. 3). Accordingly, in Series II curvatures were not measured unless it was necessary to determine whether the primary energy lay within the chosen limits.

Table II *A* and the lower curve in Fig. 3 show the dependence of *R* on primary energy. They are based entirely on the data of Series I, the energy spectrum being the solid curve in Fig. 2. The upper curve shows the data of Williams and Terroux recalculated as described in section I. The short lines at the top of the figure show the limits of the primary energy intervals.

Table I *B* and Fig. 4 showing the dependence of *R* on T_0 are obtained by combining relevant parts of the data of Series I with those of Series II. The energy spectrum of the primaries is the dotted curve in Fig. 2. In both figures the indicated errors are standard deviations due to fluctuations in the occurrence of side branches, no allowance being made for instrumental errors. In Fig. 4, the abscissæ of some points are slightly shifted to avoid overlapping.

VI. Discussion

Variation of R with T_p.—Figure 3 indicates that the large and systematic variation of *R* with T_p found by Williams and Terroux is not real. The numbers of observed particles for their 4 points are 8, 12, 20, and 14, an average of 14; in our case the numbers are 18, 24, 23, 27, 20, and 4; the average is only 19. Therefore no great weight is attached to the detailed course of either curve, but our data for the ranges 0.67–0.99 and 1.00–1.32 MEV, taken together, refute the value 2.8 found by Williams and Terroux in the same energy region; their result, indeed, may be due to a very unfavorable fluctuation.

Comparing our curves in Fig. 3 and Fig. 4, we see that exclusion of branches with energies between 12 and 20 EKV results in lower *R* values. This strongly indicates that the discrepancy between theory and our curve in Fig. 3 is partly due to the inclusion of branches down to 3 mm in length. The same conclusion applies *a fortiori* to the Williams-Terroux curve in Fig. 3, which is based

TABLE II

Series	Meters of Primary Track	Range of T_p (MEV)	Range of T (EKV)	Branches Observed	Branches Calculated	R
A; I alone	105.6	0.67–0.99	12 up	18	11.1	1.62±.38
	105.6	1.00–1.32	12 up	24	16.5	1.45±.30
	105.6	1.33–1.65	12 up	23	13.5	1.70±.35
	105.6	1.66–1.98	12 up	27	19.4	1.39±.27
	105.6	1.99–2.31	12 up	20	17.4	1.15±.26
	105.6	2.32–2.64	12 up	4	5.5	0.73±.37
		0.67–2.64	12 up	116	83.4	1.39±.13
B; relevant parts of I and all of II	190.2	0.67–2.64	20 up	107	88.3	1.21±.12
	190.2	0.67–2.64	30 up	63	56.4	1.12±.14
	190.2	0.67–2.64	40 up	35	40.9	0.86±.15

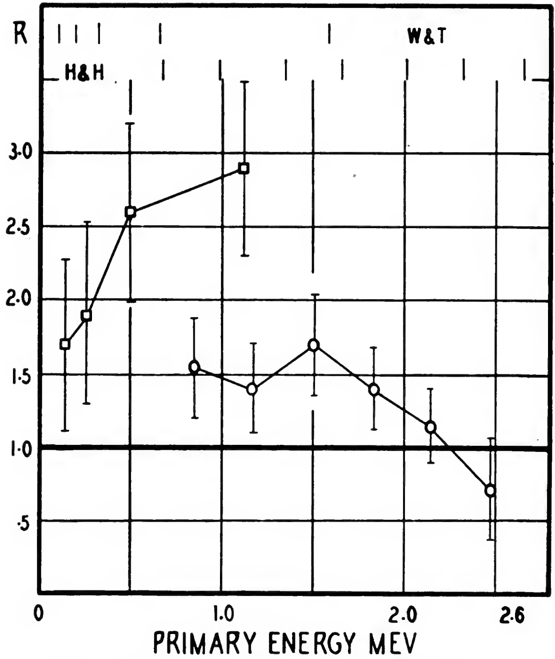


FIG. 3. Ratio R of the observed and Möller cross sections, as a function of T_p. The secondaries have all energies from 12 EKV up. See Table II *A*.

on secondaries with energies down to 10 EKV. We did not care to pursue the matter further, for curvature measurements are time-consuming, and it appeared that more useful information could be obtained by concentrating on the dependence of the cross-section on T_0.

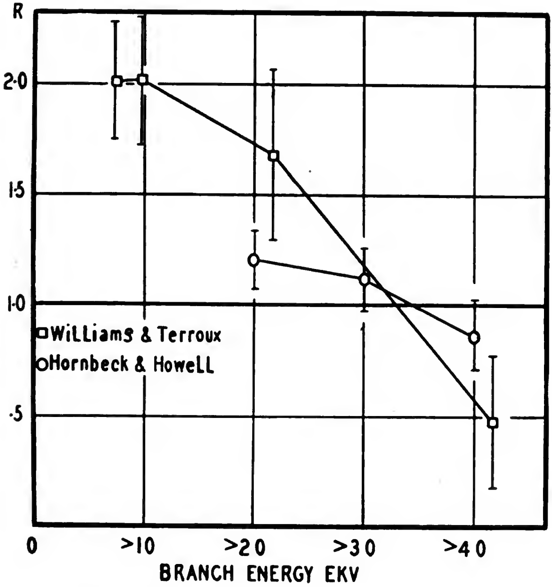


FIG. 4. Ratio R of the observed and Möller cross sections as a function of T_0, the lower limit of secondary energy. Curve of Howell and Hornbeck is based on the data of both Series I and Series II; see Table II B.

Variation of R with T_0.—Referring to Fig. 4, we first call attention to the fact that the data for each point P include the data on which points to the right of P are based. This method of analysis has some advantage over that in which one considers the values of R for successive ranges ΔT of secondary energy, because the statistical error decreases as we pass toward the left. If this were the only important consideration, the points for 7.5 and 10 EKV on the curve plotted from the Williams-Terroux data should be the most reliable ones on that curve. We note, however, that they depart farthest from theory, and in view of the relative flatness of our own curve, we believe that these points are seriously in error because of the inclusion of very short branches. With the conditions under which Williams and Terroux worked (2/3 normal pressure, room temperature) 7.5 and 10 EKV secondary tracks are 1.6 mm and 2.9 mm in length, respectively. The other two points depend on 18 and 3 events respectively; of course the latter is included here only for the sake of completeness.

Our own curve indicates the essential correctness of Möller's formula, but since the point at 20 EKV, presumably the most reliable, lies at the ordinate 1.21, we must discuss the influence of the more important errors. There are two systematic errors which cannot be eliminated, and which are in the right direction to explain this. It is difficult to correct for them with accuracy, so we have preferred to present the data as they stand, mentioning these corrections subsequently. They are:

1. Simulation of electron-electron collisions through accidental production of slow photoelectrons and recoil electrons close to a primary track. The correction is probably not more than 0.8 per cent (Sec. IV).

2. Inclusion of preëxpansion tracks in spite of precautions taken to minimize such inclusion. Before expansion the gas density is about 13 per cent greater than afterwards. The number of secondaries per meter of preëxpansion track is therefore 13 per cent too large, but we do not believe that more than 10 to 15 per cent of the measured tracks were formed before expansion. This judgment is based on comparison of the width of the widest tracks employed with that of the sharpest ones obtained (see criterion 5). Even if the fraction were as high as 25 per cent, the correction to R would be only about 3 per cent.

We come now to random errors, and errors of unknown sign, which affect either the observed or the calculated cross-section. The error in T_0 due to inaccuracies in the range-energy curve can scarcely be more than 4 per cent, since $T \sim r^{1/2}$. The effect of straggling on the average range of a group of n tracks can be estimated only roughly. The relative standard deviation of the range is of the order $26/n^{1/2}$ per cent,[21] for an energy of the order 20,000 EV. Thus if n is 35 the standard deviation of the average of the energies is about 4 per cent. It must be remembered that in this study we did not have to make range measurements unless the range was close to one of the limiting values employed (such as 1.8 cm, corresponding to 30 EKV). *Random* personal errors in measuring ranges are unimportant; we believe the length of 8 mm branches cannot be systematically in error by more than 0.6 mm, corresponding to an energy error of about 4 per cent, of unknown sign.

---

[21] Williams, *Proc. Roy. Soc.*, 130, 310 (1933); especially p. 317.

To summarize, we have discussed two systematic errors which are certainly in such a direction as to increase R. The correction for them cannot be more than 3.8 per cent and it is probable it is not more than 2.8 per cent. The random errors, and those of unknown sign, compounded in the usual way, give a resultant of 7 per cent. For our curve in Fig. 4, then, consideration of the more significant sources of error leads to the following values of R (using 2.8 per cent for the reduction due to the two known systematic errors):

T_0 in EVK	R
20	1.18 ± 0.14
30	1.09 ± 0.16
40	0.84 ± 0.13

These data indicate a slight excess of observed branch-frequency over the requirements of Möller's formula, because the first entry is presumably the most significant one. The average discrepancy with theory is almost certainly less than 18 per cent in the domains of primary and secondary energy which we have covered. Champion's results for the lower energy range 0.4–0.9 MEV deviate from the formula in the opposite direction by 8.4 per cent. Both investigations support each other in demonstrating the essential correctness of Möller's theory of secondary energy distribution in the domain of primary energy from 0.4 to 2.6 MEV. This statement applies of course to values of T/T_p in the limited range from about 0.02 to 0.1.

Thanks are due to The American Philosophical Society for a grant to Dr. Arthur E. Ruark which made this work possible. We are also grateful to Dr. Creighton C. Jones for permission to use the photographs upon which the measurements were made; and to Dr. Ruark for suggestion of the problem and for his helpful advice throughout the investigation.[22]

---

[22] A letter from Professor E. J. Williams informs us that unpublished work performed by himself and Mr. Cameron in 1933, using Ra E beta particles, resulted in R-values of about 1.2; further details are not now available.

AUTONOMOUS VERSUS REFLEXOGENOUS ACTIVITY OF THE CENTRAL NERVOUS SYSTEM [1]

PAUL WEISS

University of Chicago [2]

(*Read November 1940*)

ABSTRACT

Experiments are reported which have produced conclusive evidence in favor of autonomous, as against reflexogenous, origin of central nervous function. Both the drive (automatism) and the pattern (coordination) of central activity are fundamentally intrinsic properties of the centers.

1. Differentiated spinal cord of amphibian larvæ was transplanted under the back skin and connected with a transplanted limb as motor effector. Such "deplanted" isolated centers, while undergoing structural disorganization, exhibit recurrent seizures of rhythmic spontaneous activity for many months. Most parts of the nervous system seem to possess rhythmic automaticity, although this remains largely latent so long as the part is connected with, and subordinated to, other centers. Reflex activity appears in the deplanted centers only secondarily, after the endogenous activity has been going on for some time.

2. The fact that the central discharge patterns of coordination, as they self-differentiate, are laid down in terms of sequences of individual muscle contractions with no regard to the effectiveness or ineffectiveness of the contractions for the animal as a whole, is proved by the development of reversed locomotion in animals of which the legs have been reversed in the prefunctional stage. No adjustments are ever observed in amphibians.

3. The fact that the basic self-differentiated patterns of coordination are unmodifiable even in mammals, is proved by the permanently reversed functioning of the legs of rats after tendon crossing (Sperry).

In conclusion, the basic patterns of coordination are centrally preformed, do not arise by trial and error, and cannot be remodeled by experience.

THE useful operation of a machine depends on two factors, energy supply and a proper structural design to transform and distribute the energy appropriately. The operation of our nervous system likewise depends on two factors, generation of nervous activity, and a pattern for the selective distribution of that activity. Without proper devices for coordinated distribution, excitation produces mere convulsions. Conversely, without driving force the most appropriately designed nervous mechanism remains an idling plant. Consequently, when speaking of the origin of nervous activity, we must split the problem into one of the drive and one of the pattern of action.

---

[1] Paper read at the autumn meeting of the American Philosophical Society in Philadelphia.

[2] The experimental work reported in this paper has been aided by the Dr. Wallace C. and Clara A. Abbott Memorial Fund of the University of Chicago and by a grant from the Penrose Fund of the American Philosophical Society.

These are not necessarily of joint origin. The drive may come from without—be *exogenous* to the nervous system,—or from within, being *automatic*. Impulses, whether generated in one way or the other, are then distributed according to patterns which may be *inherited*, that is, devised by autonomous laws of the developing nervous system, or individually *acquired*, that is, molded by extraneous influences, through practice and experience. External incentive versus automatism, stimulus-determined versus autonomous response,—these are the perpetual issues facing the student of behavior. To decide among them, is an empirical problem and not one of logic. Therefore, we shall not enter into the academic discussion which has been carried on about the subject by physiologists, psychologists, educators, and even sociologists, but merely present some of the modest earth-bound efforts of the experimental biologist to provide those discussions with a more solid foundation of facts.

First, let us consider the *drive* of the nervous system. For a long time physiologists have been partial to the thesis that it is basically exogenous rather than automatic, reflexogenous rather than spontaneous. The organism was represented as a reflex machine with input, transformer, and output, and all nervous activity was thought to arise from a sensory stimulus or a chemical stimulant carried in the blood. The first serious breach in the pure reflex theory came with the discovery that the respiratory center continues to discharge rhythmic bursts of impulses even when completely isolated. Spontaneous discharges of this order were obtained by *Adrian* and coworkers (*Adrian* 1932) and by *Bronk* and *Ferguson* (1935) from the respiratory center of vertebrates, and by *Adrian* and by *Prosser* (1936) from the ganglia of some arthropods. Lately, the periodical fluctuations of electrical potential in the brain cortex, so-called brain waves, have been advanced as evidence of the spontaneous activity of this highest nerve center.[3] Yet, these cases could still be considered as exceptions rather than the rule, so long as all other nerve tissue could be thought of as devoid of spontaneous activity.

---

[3] The fact that the amphibian embryo displays motor activities before the sensory system is developed (*Coghill* 1929) furnishes a strong argument in favor of central automatism; however, occasional doubts have been raised regarding the conclusiveness of this argument on the grounds that the possibility of irritation of the centers by substances carried in the blood or other body fluids was not excluded.

However, this thought can no longer be maintained in the light of the experiments which I am about to report, and which have shown that the capacity for spontaneous, rhythmic, activity is indigenous to most central nervous substance, although it takes special measures to bring it out. These measures consist of allowing the centers to undergo some degree of destruction and degradation. This is achieved by a method which we may call "*deplantation.*"[4] Excised fragments of central nervous system are embedded in a well-vascularized, but otherwise indifferent, tissue. The gelatinous connective tissue of the dorsal fin of amphibians is excellently suited for this purpose (Fig. 1). A bit of brain or spinal cord deplanted into this tissue suffers serious disorganization and loses part of its cell content by degeneration. The remaining part, however, survives and becomes functional. To assay its activities, we provide the deplanted center with the most natural and adequate detector, namely, a group of muscles: We transplant a limb into its vicinity, taking care that nerve connections can be established between the two grafts. Thus, we produce an isolated complex, consisting of a deranged nerve center, a limb, and random nerve connections between them, the whole unit supplied by the common blood stream, but otherwise independent from the host animal (*P. Weiss* 1940).

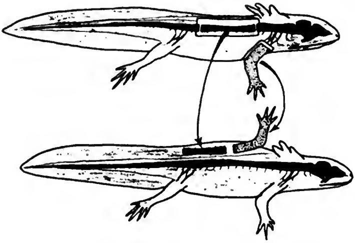


FIG. 1. Deplantation of a fragment of spinal cord (black) and transplantation of a limb (stippled) to the dorsal fin of a salamander larva. Within a few weeks nerves from the deplant grow into the limb effecting functional connections with its muscles and skin.

No sooner have the nerve fibers from the deplanted center united with the muscles of the limb graft, than the latter begins to register almost incessant *spontaneous activity.* Without any external stimulus, in a perfectly quiescent host body, the grafted center discharges into the limb in intermittent spells which range from continuous arhythmic fibrillations to strong synchronized pulses of a fairly regular rhythm of from 20 to 90 per minute. Frequently a seizure begins with weak irregular twitches and gradually builds up in strength to violent and more synchronized activity during which the limb muscles beat in unison at fairly regular intervals. The seizures come and go for weeks, continuing uninterrupted for hours, and the whole phenomenon has been seen to last for as much as 5 months, which was as long as the animals were kept. The fact that the observed activity really originated within the deplanted centers, was, of course, established beyond doubt by a variety of checks.

---

[4] *Przibram* (1930) defines "deplantation" as any "insertion of a removed part in a different location." In the more restricted sense in which we propose to use the term, however, it connotes a marked degradation of organization suffered by a part which has been transplanted into an atypical location.

Reflex responses to touch also appear, but not until many weeks after the onset of spontaneous activity. Hence, the preceding pulsating activity is definitely of automatic and not of reflex origin. It has been obtained thus far from the following variety of grafts (*P. Weiss* 1941): spinal cord from any region and of any length; medulla oblongata; spinal cord slashed or minced and then implanted as a single lump; spinal cord from larvæ as well as from adults; fragments from purely ventral or purely dorsal sectors of the spinal cord, which is particularly remarkable in view of the fact that the dorsal half contains no motor cells, and the ventral half no sensory elements.

The constancy of the phenomenon under such a variety of conditions intimates that we are dealing with a fundamental capacity of all nerve substance which is manifested as soon as the more specialized superstructure of anatomical and functional relations, present in the normal developed nervous system, has broken down. The close connection between derangement and the emergence of automatic seizures can be directly demonstrated (*P. Weiss* 1941). If a spinal limb center and the corresponding limb are deplanted together under careful preservation of their nerve connections (Fig. 2), orderly reflexes are obtained immediately after the operation and during the following days. Gradually, however, the reflex responses become cruder, protracted, and generally disorganized. Significantly, *pari passu* with this deterioration, spontaneous activity makes its appearance.

Supported by histological evidence, we may consider our deplants as neurone pools in which the typical organization has given way to a random make-up. Such a random neurone pool, however, exhibits many of the functional properties of centers which we have been used to attribute to a definite and minutely arranged neurone architecture. Its discharges are rhythmic, and the participating elements tend to become synchronized with increasing intensity of the reaction. A certain analogy to the heart beat is obvious, and one wonders whether a pacemaker mechanism or some mutual induction among the elements is responsible for the synchronization; but this is for the future to decide. We are left with the impression that something is constantly being built up in the centers which, upon reaching a certain threshold value, gives rise to a discharge, whereupon a new cycle begins. Again, whether that something consists of an electrostatic field, or metabolites, or something else, only the future will tell with certainty. Our preparations lend themselves readily to a further investigation of these problems.

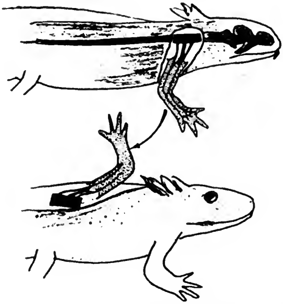
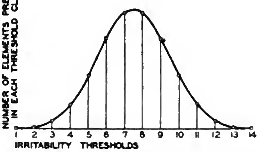


FIG. 2. **Deplantation of limb segments of spinal cord with nerves and innervated limb remaining attached.**

The reflexes obtained from our neurone pools show some typical attributes of normal reflex action, namely, graded response, spatial summation, temporal summation, and fatigue. In addition to releasing motor responses, afferent impulses also seem to lower the excitability of the whole pool, which often results in a burst of spontaneous activity trailing the reflex.

In view of the random character of our pools, their functional properties will have to be explained on purely statistical assumptions. If we assume that the thresholds of the cell population are distributed according to mere chance, the numbers of cells in each threshold class, plotted against thresholds, would form a normal distribution curve (Fig. 3). Since an afferent input of given strength will activate all elements of the corresponding and lower thresholds, the number of discharging elements is expressed by the area of the binomial function up to that threshold value. Since the curve representing the integral of the binomial function is in the major part of its course sufficiently close to a straight line (Fig. 4), we realize that in a random pool of neurones there will be some proportionality between sensory input and response.[5] The implications of this thought for the physiology of the normal nervous system are obvious, but have not yet been followed up in detail.



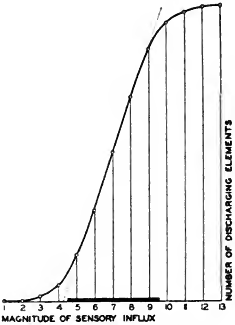


Fig. 3. Normal distribution of the units of a neurone pool according to threshold classes, on the assumption that constitutional threshold variations are of random character.

Fig. 4. Area enclosed by the curve of Fig. 3 (in arbitrary units) plotted against the individual threshold values. This curve follows within its middle range (marked by a black bar) a nearly linear course. Since this range includes almost 90 per cent of all elements present, it follows that every stepping up of the sensory effect, either by increase of the peripheral stimulus or by spatial or temporal summation, produces a corresponding increase in the number of discharging efferent units.

---

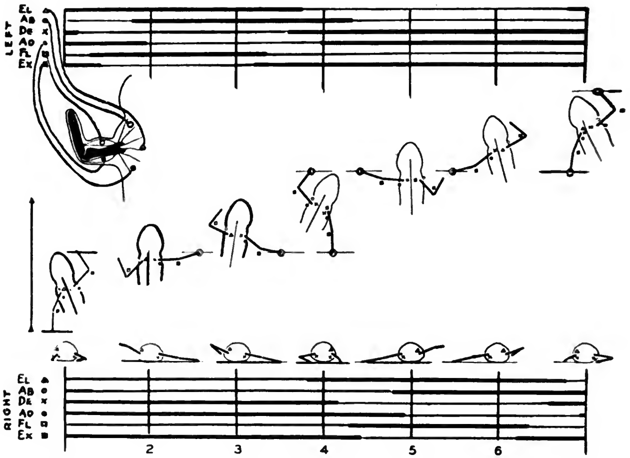
[5] What the observations have thus far established, is only an estimated gross proportionality between stimulus and effect; but whether this proportionality follows a linear or some other function has not been ascertained and will remain unknown until quantitative tests have been applied. We merely intend to show that approximate linearity would bear out the randomness of the neurone distribution in the pool.

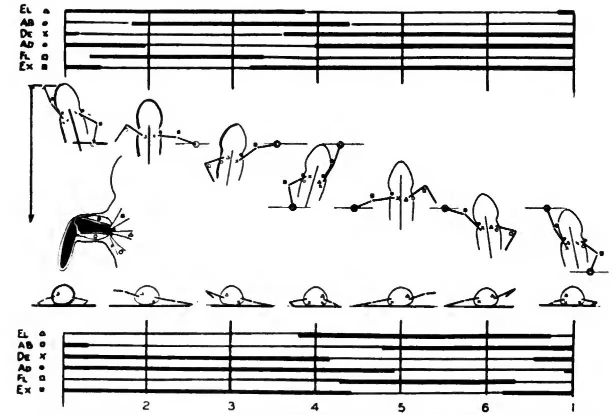
The intensity of spontaneous activity fluctuates with the composition of the blood of the host animal. Changes in the blood produced by light exercise of the body cause a marked increase in the excitability of the deplants, which expresses itself in increased spontaneous activity and lower reflex thresholds. Whether anoxemia, acidity, carbondioxide, or metabolites are involved, has not yet been determined. This sensitizing effect is often sufficient to start a temporarily quiescent deplant off on a spell of activity. Observations like these suggest that the cyclic process underlying the rhythmic spontaneous discharges is actually going on in the neurone pool all the time, whether a discharge appears or not. An effective discharge will result only if the process reaches a critical threshold intensity. Otherwise it merely conditions the nerve cells so that they can go off under such additional stimuli as may come in from other sources, provided the total sum adds up to threshold intensity.

Viewed in this light, the endogenous pulse of a nerve cell becomes of fundamental importance for nervous activity, even when it does not register immediately as a visible motor effect. And instead of regarding spontaneous activity as an exceptional manifestation of cortex, respiratory center, and insect ganglia, we may have to concede it to all nerve substance. Then, of course, the question arises as to what causes it to remain latent in most of the normal centers. We do not know the answer, but our experiments suggest that the repressing effect is somehow tied up with the normal organization of the centers, and disappears with the latter. In this connection we may point to the striking resemblance between the seizures of our deranged deplants and epileptiform seizures in human patients. It will be interesting to examine whether this resemblance is purely superficial or really pertinent.

Let us now turn to the second point of our program—the *patterns* of nervous activity. Are they of central or of extraneous origin? Recent experiments have given an incontrovertible answer. The patterns underlying the basic acts of behavior of an animal are developed by *self-differentiation* within the centers, without experience, and, indeed, sometimes in spite of experience.

Coordinated locomotion requires that the muscles be engaged in a definite chronological order. The correct timing and dosing of the individual contractions is all that distinguishes the harmony





Figs. 5 and 6. Diagrams indicating the "scores" of muscle action in the fore-limbs during a single step of a walking salamander, as reconstructed from motion pictures.

Fig. 5. Normal animal.

Fig. 6. Experimental animal with reversed fore limbs. Only six main muscle groups are recorded: elevators, abductors, depressors, and adductors of the upper arm;

of the gait from the discord of a convulsion. The muscles are played by the centers according to a definite "score," like an orchestra (Fig. 5). Now, previous experiments (*P. Weiss* 1936) have shown that the muscle scores for locomotion are laid down in the centers in terms of the individual muscles, muscle for muscle, and that these scores, at least in the lower vertebrates, are fixed and inaccessible to reconditioning by experience. If, for instance, we disarrange the muscles so as to make their play yield incongruous effects, the centers, nevertheless, continue to call them into action according to the old score, as before; and they never learn to do better. The most extreme case is illustrated by the following experiment on salamanders (*P. Weiss* 1937). A limb is replaced by its counterpart from the opposite side. This amounts to reversing the position of the flexor and extensor muscles with regard to the body as a whole (Fig. 6). The central nervous system, however, persists in sending out the impulse pattern in the old order meant for the normal arrangement, and, consequently, such animals with reversed limbs, when bent on forward locomotion, actually move backward, and when intending to retreat, actually advance (Fig. 6). In spite of their predicament, the animals have never been able to adjust their coordination to the new situation. There is no trace of reeducation, and the scores of coordination, once established, prove unmodifiable.

Since coordination cannot be relearnt, it seems unlikely that it should have been learnt in the first place. Nevertheless, a crucial test had to be made, and this was done by reversing the musculature in the early larval stage, before there had been any function. When such animals later take up ambulation, the limbs again move in reverse, and do so throughout life. In other words, the central nervous system discharges according to the score for a normal unreversed limb, although it has never had a chance actually to operate a normal limb, as none is present.

This is definite proof that the basic patterns of coordination self-differentiate within the central nervous system. They differ-

---

flexors and extensors of the fore-arm. The position of these muscles is indicated in the inset. The record for the left limb is on top and that for the right limb at the bottom of each diagram. In the center seven phases of the actual movement of the animal are presented in dorsal and rear views; the points in which the hand is put down on the ground, and which serve as pivotal points for the body during the subsequent phase of movement, are marked by circles.—The "scores" are identical for the normal (Fig. 5) and reversed (Fig. 6) limbs, but, in the latter case, result in a backing movement.

entiate into what would be the correct score for a normal limb, regardless of whether or not the movements which they produce will actually be of service to the animal. This does not preclude learning by experience in lower vertebrates altogether. It merely specifies that the learning process must avail itself of the basic preformed patterns of performance; learning may be able to facilitate, suppress, dissociate or rearrange combinations of *existing* patterns, but it cannot break the individual patterns down and reconstruct or remodel them to suit emergency situations.

In view of these results on lower vertebrates, a re-investigation of the situation in higher mammals became imperative; for "natura non facit saltum." A student of mine, Mr. Sperry, was charged with exploring in the rat the effects which crossing of the tendons of antagonistic foot muscles would have on locomotion (*Sperry* 1940). Each muscle pulling on the wrong side of the skeleton (Fig. 7), at first, of course, produced effects, which were the reverse from what the animal intended. But according to the literature a gradual recovery of normal function should have been expected. No such readjustment, however, occurred even under most crucial training conditions, the muscles maintaining their reversed play (Fig. 8) throughout life. Basic coordination patterns are as fixed in the rat as they are in amphibians, and experience can have had

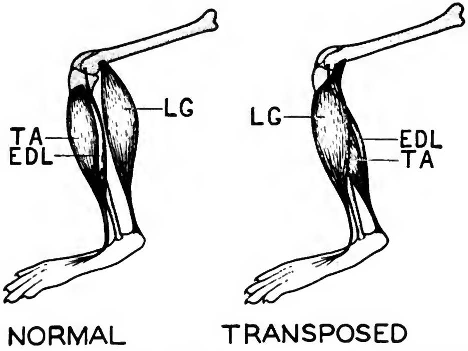


FIG. 7. Transposition of the tendons of the leg muscles in the rat so as to reverse the mechanical action of these muscles (after Sperry). *TA*, anterior tibial muscle; *EDL*, long extensor of the digits; *LG*, lateral gastrocnemius. (All other muscles removed.)

no part in their establishment. The same is presumably true of higher mammals.

But how about man? Do we not know from innumerable experiences of surgeons that a patient whose paralyzed muscles have been substituted by healthy muscles of different function, gradually learns to put the new combination of muscles to good use in walking and other performances? Jointly with our Orthopedics Department, we are conducting at present a thorough reinvestigation of this whole problem. But even on the basis of our present knowledge, it seems safe to predict, that whenever a patient learns to use a muscle in a different sequence than that provided for in the inherited pattern of gait, this learning act is on the same level as the learning of trick movements or other skills; in other words, it is a *new* performance, not a remodeled old one. Voluntary effort of the cortex can do these things which lower vertebrates cannot do. While even conscious effort cannot recondition the basic inherited patterns of coordination, it can learn to supersede them by novel ones, which will last so long as the cortex acts, obscuring, but not abolishing, the lower autonomous patterns. Apparently, mammals have acquired the ability to engage muscles in arbitrary combinations only during fairly recent stages of their evolutionary course. This fact is significantly reflected in the increasing prominence which the pyramidal (cortico-spinal) system has gained in higher mammals: through this tract the cortex can activate spinal motor neurones directly by a short cut around the hierarchy of subcortical mechanisms which operate the inherited patterns.

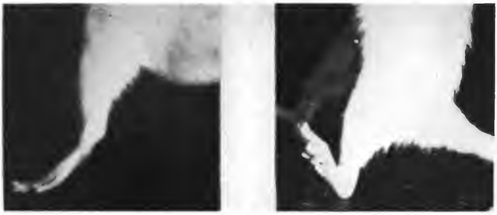


FIG. 8. Foot action in the extension phase of a movement. (After Sperry.) (*a*) Rat the tendons of which were severed and resutured without crossing; plantar flexion of foot. (*b*) Rat with crossed tendons; reversal of foot movement, *i.e.*, dorsal flexion.

Coordination *by experience*, therefore, is a secondary, and highly specialized, affair. All primary coordination exists *by predesign*, established in the pre-functional phase of neurogenesis, independent of, and unresponsive to, experience.

In summarizing, we recognize that the nervous system, far from playing the passive role which at times has been attributed to it, provides part of its drive and part of its patterns of action from its own inherited resources. *Automatism*, though manifesting itself under normal conditions only in certain kinds of centers (respiratory center, arthropod ganglia), may yet exist in all centers as a conditioning mechanism, disposing for discharge, rather than actually discharging, central neurones. And *autonomy* must be conceded at least to the basic patterns of performance of all animals from the lowest to the highest.

BIBLIOGRAPHY

ADRIAN, E. D. 1932. *The Mechanism of Nervous Action.* Johnson Foundation Lectures. Philadelphia.

BRONK, D. W. AND L. K. FERGUSON. 1935. The nervous control of intercostal respiration. *Am. J. Physiol.*, **110**, p. 700.

COGHILL, G. E. 1929. *Anatomy and the Problem of Behavior.* Cambridge University Press.

PROSSER, C. L. 1936. Rhythmic activity in isolated nerve centers. *Cold Spring Harbor Symposia on Quantitative Biology*, **4**, p. 339.

PRZIBRAM, HANS. 1930. *Experimental-Zoologie*, vol. 7: Zootechniken. Leipzig and Wien.

SPERRY, R. W. 1940. The functional results of muscle transposition in the hind limb of the rat. *J. Comp. Neur.*, **73**, p. 379.

WEISS, PAUL. 1936. Selectivity controlling the central-peripheral relations in the nervous system. *Biol. Reviews*, **11**, p. 494.

——. 1937. Further experimental investigations on the phenomenon of homologous response in transplanted amphibian limbs. IV. Reverse locomotion after the interchange of right and left limbs. *J. Comp. Neur.*, **67**, p. 269.

——. 1940. Functional properties of isolated spinal cord grafts in larval amphibians. *Proc. Soc. Exp. Biol. and Med.*, **44**, p. 350.

——. 1941. Further experiments with deplanted and deranged nerve centers in amphibians. *Proc. Soc. Exp. Biol. and Med.*, vol. **46**, p. 14.

RESPONSIVE BONE *

C. B. DAVENPORT

(Read Nov. 22, 1940)

ABSTRACT

Between dead bone and muscle there is clearly the greatest possible contrast in responsiveness to stimuli. However, it has long been known that the trabeculæ of the upper end of the femur are directed in lines so as to resist the weight of the body most efficiently with the least expenditure of material. It has been shown that if after a break the bone is set badly new trabeculæ are formed to meet the new direction of the thrusts upon the bone. A study of the X-rays of the bones in the foot afford new and striking evidence. The heel bone of a normal, walking person is shown to contain dark lines, rods or trabeculæ, the direction of which is in line with that of thrusts and pulls upon the bone. In a 20-year-old girl who, through birth injuries, has never walked, the heel bone is still formed roughly in the normal shape but is without the properly directed trabeculæ. In the heel bone of a girl of 20 years, who walked until 10 years of age, but who has been bedridden since, the heel bone has a nearly normal form, but only a few directed trabeculæ.

The conclusion seems to be justified that the bone-forming cells respond to directive thrusts and pulls made upon normally functioning bones by forming the trabeculæ in adaptative positions. Thus the bone cells are as truly responsive as muscle cells. The difference between the two kinds of cells is this: The muscle cells are constructed to respond to stimuli by contracting and the bone cells to respond by supporting, or rather, building structures that shall resist blows or pulls. The bony cells are as truly responsive as the muscle cells but respond to a different group of stimuli and in a different way.

Popularly considered, there is the greatest possible contrast in responsiveness between muscle and bone. Under the stimulus of an electric potential, whether applied outside the body or from a nerve in the living active individual, the muscle contracts. The response, if any, in the bone is less obvious. It has, however, for 70 years been known that during its development the thigh bone shows in its structure evidence of something like adaptive response (Meyer, 1867; Wolff, 1870). The weight of the body (in extreme cases 300 pounds) is carried by the head of the femur and transferred through the neck, 3 cm. in length, to the shaft of the femur. If the femur were solid hard bone it could easily carry this weight, but at the expense of a large amount of material and increase of the weight of the

---

* Acknowledgment is made of assistance from the Penrose Fund of the American Philosophical Society in preparation of this paper; also from the Carnegie Institution of Washington and Department of Mental Hygiene, State of New York.

bone. Actually, the interior of the head and neck of the femur is made up of a trestle work consisting on the one hand of plates which transfer the pressure on top of the head to the shaft of the femur and, secondly, of plates or rods which tie the first set to the opposite face of the femur, as though better to distribute the pressure. A third set of plates strengthens the second. It was long ago shown by an engineer (Culmann) that these plates are laid down in accordance with the engineering rules of graphical statics, Fig. 1. The problem has then arisen: In what manner are the bone-forming cells guided in forming these rods in the appropriate positions?

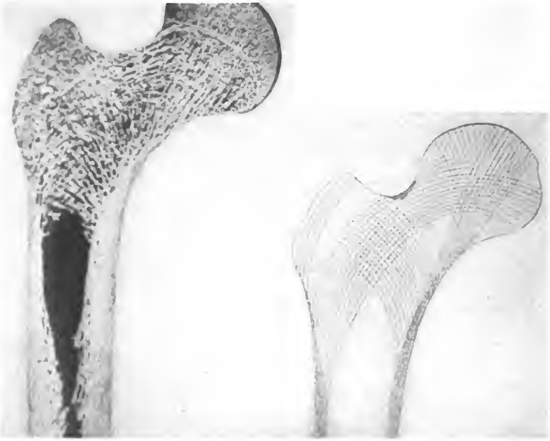


Fig. 1. Left: Median section of the upper end of the femur showing lines of bony rods in spongy tissue. Right: Statical diagram of lines of force when weight is applied at the top of the head of the femur, as drawn by an engineer.

It has also been shown that if the shaft of the femur of a normal person is broken and subsequently the parts are badly knit together so that the shaft is bent at an angle then the direction of the plates in the head and neck is redistributed. Evidently the old plates are torn down and new ones, better suited to meet the statical conditions now imposed, are formed.

Roux (1865) classified the above described phenomenon under the head of functional adaptation without going further into the interpretation.

Additional light upon the interpretation of this condition is thrown by observations I have been able to make upon the feet of some living children in the third decade of life. These observations are a sort of by-product of a study financed in part by the Penrose Fund of the American Philosophical Society.

Figure 2 shows the radiograph of a normal walking person. The heel bone shows numerous dark lines which are the bony rods that continue the lines of pull of certain muscles and tendons and the blows that the heel bone receives in walking. At the upper part of the heel bone is attached, just below the notch, the muscle known as quadratus plantæ. This is the muscle which serves to maintain the bony arch of the foot. It is the muscle that becomes flacid when not used so that its function has to be replaced in some unfortunate persons by a steel arch in the shoe. From the point of attachment bony rods are seen running across the heel bone to support the pull of the muscle. Below the quadratus muscle is the long plantar ligament which ties together the arch of the foot fore and aft to

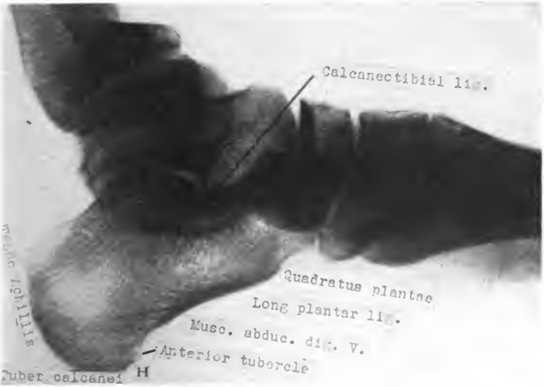


FIG. 2. Heel bone of a normal walking person. Heel bone (*H*) shows numerous (dark) bony rods that resist the blows and pulls upon it. Muscles and ligaments lie approximately in the position of the labels or are attached as indicated.

prevent its extreme distension. It is attached along a centimeter or two of the heel bone and the bony plates here are most numerous and strongest of all. Still below is the muscle that abducts the toes and in this case also the attachment is indicated by bony plates. The area of attachment of this muscle reaches down to the so-called anterior tubercle of the heel bone to which the front end of the walking pad is attached. From this anterior tubercle to the posterior tubercle is the region which receives the blows of the pavement and from this region bony rods run forward to attach to the largest bone of the tarsus, the talus. Into the hind tubercle is inserted the tendon of Achilles, which runs upward into a prominent calf muscle that lifts the heel in walking. From the area of insertion of this tendon bony rods run downward and forward toward the anterior tubercle. Finally there is a ligament which unites the heel bone with the tibia, or shin bone, and the stresses of this ligament are resisted by a complicated set of rods irradiating from the center of the thrust.

Figure 3 shows radiographs of a child of 20 years who has never walked.[1] The history as obtained by personal investigation is conclusive on this point. One notes that the heel bone is formed. This much has an hereditary or genetical basis. One notes too that the anterior tubercle is formed and that there is a hollow above it to which the ligaments and muscles are attached. The shadows of such can be seen in the radiograph. Also the posterior tubercle is formed and no doubt the tendon

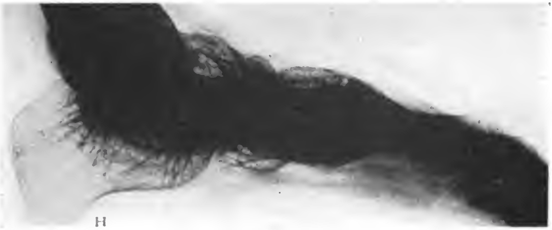


FIG. 3. Heel bone (*H*) of a child of 20 years who has never walked. Bone imperfectly shaped; has a few irregular rods.

---

[1] Weidenreich (1922) refers to a somewhat similar case.

of Achilles is there also. This much has apparently been developed independently of the functioning of the bone. If, however, one looks at the structure of the bone one finds first that the rectangular system of resisting rods is hardly represented or, at best, only by fine lines forming an irregular web, but not in the line of direction of the pulls as found in walking children. One may say that the regular rods that are normally present are formed only in response to stresses; no stresses no rods.

Figure 4 shows a radiograph of the heel bone of a girl of 20 years who walked until she was 10 years old but has been bedridden since. The bone has a nearly normal outer form. The tubercles are well developed and the embayment above the anterior tubercle affords ample room for the ligaments and muscles that are attached there. One may say that the genetical factors involved in the production of the normal form of the bone have been assisted by the temporary act of walking.

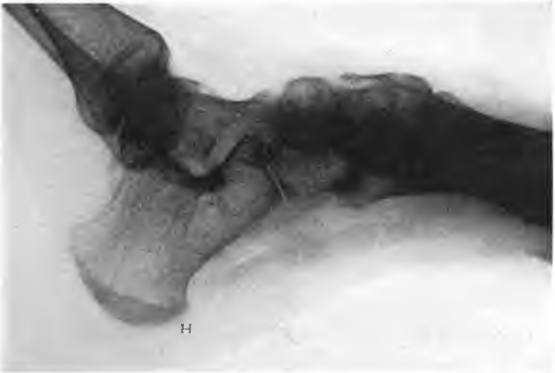


FIG. 4. Heel bone (*H*) of a girl of 20 years who walked until 10 years of age but has been bedridden since. The bone has a nearly normal outer form but only a few rods.

However an examination of the bony rods in the interior of the bone show that they are feebly developed. There are a few rods which run from the attachment of the plantar muscles and ligaments and there are some vertical rods which receive the pull of the calcaneo-fibular ligament, and a few feeble ones that would resist the pressure of the heel on the pavement. But, on the whole, the system of bony rods is vastly reduced. We conclude then that a system of rods once formed in response to the functioning of the heel and foot has been largely resorbed in the course of years during which the foot has not functioned.

The conclusion from these radiographs, a conclusion which is supported by a considerable number of others, is that in the formation of bony plates the bone is responsive to the stimulus of stress. Just as the muscle responds to stimulus by contraction so the amœboid bone-forming cells of the heel bone respond by forming rods in appropriate position to resist the stresses applied to the surface of the bone. The business of the muscle is to contract when called upon to do so; the business of the bone-forming cells is to afford support when called upon to do so. Our false idea of the rigidity of bone is based upon the dried specimens that we see in museums. These are not, physiologically speaking, bones, but only the skeletons of bones.

Just how the bone-forming cells are capable of responding to the stress of the muscles by forming rods is obscure, just as the mechanism by which an electric potential applied to a muscle leads to contraction is obscure. Possibly pressure applied to the surface of a honey-combed viscous mass is transmitted in pressure lines through that mass. Apparently the amœboid cells in the living bone respond to these lines of pressure just as a living amœba responds to contact with threads of algæ or glass hairs in the medium in which it is moving, or as fish orient themselves in the flowing stream.

So we reach the conclusion that bone is just as responsive to stimuli in its way as muscle is in its way.

LITERATURE CITED

MEYER, H. v. 1867. *Arch. Anat. Physiol.*
ROUX, W. 1885. *Arch. Anat. Phys.*, Anat. Abt.
WEIDENREICH, F. 1922. *Arch. f. Entw.-Mech. d. Organ.*
WOLFF, J. 1870. *Virchow's Arch. Path. Anat. Physiol.*

NATURAL SELECTION BEFORE THE "ORIGIN OF SPECIES"

CONWAY ZIRKLE

Abstract

The history of the concept of natural selection has generally been traced back through the personal development of Charles Darwin to Thomas Malthus, whose *Essay on the Principles of Population* gave Darwin the clue which led him to formulate the doctrine. Actually the conception of natural selection is very old, although originally it was not used to explain the origin of new species (evolution) but to account for the existence of adaptation. The survival of the fit organism, of course, implies the survival of fitness itself, and thus natural selection can serve as an alternative explanation of those facts which are generally cited as evidences of teleology. Natural selection was used for this purpose by Empedocles (400 B.C.), Lucretius (99–55 B.C), Diderot (1749), Maupertius (1756), and Geoffrey St. Hillaire (1833); but it was specifically rejected in favor of teleology by Aristotle (384–321 B.C.), Lactantius (260–340 A.D.), St. Albertus Magnus (1236), and Whewell (1833). Natural selection was used to explain organic evolution by Wells (1813), Matthews (1831), Darwin (1858), and Wallace (1858).

As an explanation of evolution, natural selection involves a number of distinct though subordinate propositions, such as the existence of heritable variations, of population pressure, of a struggle for existence and the consequent survival of the fit or better adapted. A number of philosophers and naturalists recognized the validity of one or more of these propositions without however, gaining any clear conception of the implications of the whole doctrine. One such component, population pressure, was described by Hale (1677), Buffon (1751), Benjamin Franklin (1751), Bonnet (1764), Monboddo (1773), Herder (1784), Smellie (1790), Malthus (1798), Prichard (1808), Wells (1813), Matthews (1831), De Candolle (1833), Lyell (1833), Geoffrey St. Hillaire (1833), and Spencer (1852). The struggle for existence was described by al-Jâhiz (9th cent.), Hobbes (1651), Hale (1677), Buffon (1751), Monboddo (1773), Kant (1775), Herder (1784), Smellie (1790), Erasmus Darwin, (1794) Wells (1813), De Candolle (1832), Lyell (1833), and Spencer (1852). Several eighteenth and nineteenth century scientists almost grasped the full significance of natural selection but just failed to recognize all of its implications. Among these were Rousseau (1749), Prichard (1808, 1826), Lawrence (1819), Geoffrey St. Hillaire (1833), Herbert (1837), Spencer (1852), and Naudin (1852).

BEFORE the end of his life Charles Darwin realized that he had become an important figure in nineteenth century biology and so, in spite of his innate modesty, he answered frankly the many personal questions his friends asked. His contemporaries also realized his significance and they preserved a great many of his letters most of which have now been made public. Darwin himself contributed an autobiographical chapter to the biography written by his son, Francis, and during the last half century he was made the subject of so many biographies, recollections and appraisals that we know more about him, perhaps, than we do about any other scientist.

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 1, APRIL, 1941

Darwin was particularly careful in describing just how he got the idea of natural selection. Through his observations of animal distribution made while he was on the four year voyage of the Beagle he was led to the conclusion that species were not constant units. He was also greatly impressed by the changes which artificial selection could make in domestic animals and he needed only to discover some selecting agency in nature to explain how new species could come into being. As a result of reading Malthus' *Essay on the Principles of Population* in 1838, he became convinced that Nature herself was such an agent. In as much as Alfred Russel Wallace, who independently discovered natural selection, was also led to its discovery by reading Malthus' *Essay*, natural selection has generally been regarded as an outgrowth of the Malthusian doctrine. Thus the most widely accepted explanation of organic evolution is considered by many to be derived from a particular set of social theories, and biology becomes indebted to the social sciences for its outstanding contribution of the nineteenth century.

Recently, this derivation of natural selection has acquired an increased prominence, particularly in the United States, where the social significance of practically everything receives attention. Cowles [1] traced the theory of natural selection from Darwin back to Malthus and forward again to Walter Bagehot. Here he pictured the concept as transferred from sociology into biology, where it was added to and modified, and then back to sociology, both sciences profiting incidentally by the cross fertilization. Sandow [2] has gone even farther. He described Darwin's intellectual millieu and concluded that Darwin was greatly influenced by the *laissez faire* philosophy of contemporary capitalism; the universal struggle for existence, which Darwin postulated, was a general principle which in human affairs normally took the form of business competition. The reason for English thinkers in the early nineteenth century playing such a dominant role in establishing the theory of evolution was that capitalism was most firmly established in England.

It should be noted, however, that those who trace the theory of natural selection back into the field of social studies or who see in natural selection an expression of a particular social philosophy are concerned almost exclusively with the emergence of the theory through the personal development of Charles Darwin. Now Darwin was exceptionally original and industrious but he rarely bothered to investigate the past history of the ideas he was so successful in developing. Any historical summary which he published would occur as often as not in the preface to a second edition where he would cite the works which had been called to his attention by his friends. When he conceived the theory of natural selection, he was ignorant of the fact that it had already been used to explain evolution, nor did he know of the work of Wells or Mathiews, when he published the *Origin of Species*. Even when he learned of population pressure from Malthus, he was apparently unaware that the idea had already been published by a number of biologists.

---

[1] Cowles, Thomas. "Malthus, Darwin and Bagehot: A Study in the Transference of a Concept." *Isis*, **26**: 341–348, 1937.

[2] Sandow, Alexander. "Social Factors in the Origin of Darwinism." *Quart. Rev. Biol.*, **13**: 315–326, 1938.

Thus in regard to natural selection, Darwin did not represent the main stream of biological thought. He was a tributary but such a brilliant and spectacular one that he obscured the main stream itself and led historians who sought the source of the idea back to Malthus. Thus the great antiquity of the concept of natural selection has generally been overlooked. It is the purpose of this paper to call attention to the growth and spread of the idea of natural selection as it occurred in biological theory.

Before we can hope to trace natural selection through its gradual and erratic development, we must break it down into its component propositions for only rarely was its whole significance understood. Many individuals who described certain portions of the theory, obviously missed the implications of the entire doctrine, so it will be well for us to have the separate parts clearly in mind when we endeavor to record the growth of the whole complex idea.

Natural Selection has been stated very simply as follows:

1. All organisms vary, no two are exactly alike.
2. More individuals are born than can exist in the available space.
3. Consequently there is a keen competition between all forms of life (Struggle for existence).
4. The weakest or the least fit are eliminated.
5. The fit survive.
6. The survivors form a basis for new variations.
7. These variations selected long enough produce new species.

If we accept these seven propositions as literally true[3] two corollaries immediately present themselves.

1. Organic evolution has occurred, that is, the innumerable species now existing were not created in their present form but have evolved from less numerous predecessors. All of the myriad forms of life now on earth are descended from a few or even a single ancestral form.
2. Adaptation or fitness itself has evolved. The survival of the fit organism of course provides for the survival of fitness.

Natural selection thus provides an alternate explanation for the facts generally cited as evidences of teleology.

One more aspect of the theory of natural selection should be mentioned before we proceed with the history of its development. The struggle for existence has been visualized so frequently as a gladiatorial combat induced by population pressure that the layman often does not realize that natural selection can operate in the total absence of organic competition. An environment unfavorable to life may eliminate some variations and select others. Certain effects of mutation pressure, *i.e.* lethal and sub-lethal genes, mutations which make for degeneration, etc., are checked and counterbalanced continually by an active natural selection even in fit species which are relatively stable. If the organism is to survive, it must be able to cope with its environment, thus not only is the fit organism selected but fitness itself is preserved and in a changing environment, a new fitness is evolved.

The first description of natural selection of which we have any evidence occurred in the great lost poem of Empedocles. Some fragments of the passage in question have been collected and while they themselves have no direct bearing on the subject, we can identify them through the later writings of Aristotle and Lucretius as all that is left of the first use of natural selection to explain adaptation.

---

[3] It is not the purpose of this paper to describe the present status of the theory of natural selection. It should be pointed out, however, that natural selection has now been developed to a point where no dilettante can be said to have a "right" to his own opinion concerning its validity. Any honest appraisal must be based on a considerable technical knowledge of the principles involved. Thanks to the fundamental work of Sewell Wright, J. B. S. Haldane and R. A. Fisher, the effects of selection upon Mendelian populations can now be evaluated mathematically. Dobzhansky's description of the role of natural selection in evolution is brief and to the point (*Genetics and the Origin of Species*, New York, 1939). We know at present that mutation pressure and the scattering of variability in small populations can produce new varieties, and that chromosome vagaries and chance hybridizations can give rise to new species, not only without the aid of natural selection but actually in opposition to it. Natural Selection remains the court of last appeal, however, and determines whether the new variety or species persists or becomes extinct.

The fragments are quoted from the translation of William E. Leonard:

62. But come! now hear how 'twas the sundered Fire
Led into life the germs, erst whelmed in night,
Of men and women, the pitied and bewailed:
For 'tis a tale that sees and knows its mark.
First rose mere lumps of earth with rude impress,
That had their shares of Water and of Warm.
These then by Fire (in upward zeal to reach
Its kindred Fire in heaven) were shot aloft,
Albeit not yet had they revealed a form
Of lovely limbs, nor yet a human cry,
Nor secret member, common to the male.
57. There budded many a head without a neck,
And arms were roaming, shoulderless and bare,
And eyes that wanted foreheads drifted by.
58. In isolation wandered every limb,
Hither and thither seeking union meet.
59. But now as God with God was mingled more,
These members fell together where they met,
And many a birth besides was then begot
In a long line of ever varied life.
60. Creatures of countless hands and trailing feet.
61. Many were born with twofold brow and breast,
Some with the face of man on bovine stock,
Some with man's form beneath a bovine head,
Mixed shapes of being with shadowed secret parts,
Sometimes like men, and sometimes woman-growths.

The nature of the missing portions of the poem, which held these fragments together, can be inferred in part by a casual reference to the "Man-faced oxen" of Empedocles by Aristotle who stated the case for natural selection only to refute it. The passage occurs in the *Physics* (II: §8). Apparently some attempts at a purely mechanistic explanation of natural phenomena had achieved sufficient prominence in the fourth century before Christ for Aristotle to consider them worthy of rebuttal. Aristotle, of course, was too much of a teleologist to be hospitable to the conception of natural selection. The following excerpt is from the translation of F. M. Cornford, London, 1929, p. 169–170:

. . . Why not say, it is asked, that Nature acts as Zeus drops the rain, not to make the corn grow, but of necessity (for the rising vapour must needs be condensed into water by the cold, and must then descend, and incidentally, when this happens, the corn grows), just as, when a man loses his corn on the threshing-floor, it did not rain on purpose to destroy the crop, but the result was merely incidental to the raining? So why should it not be the same with natural organs like the teeth? Why should it not be a coincidence that the front teeth come up with an edge, suited to dividing the food, and the back ones flat and good for grinding it, without there being any design in the matter? And so with all other organs that seem to embody a purpose. In cases where a coincidence brought about such a combination as might have been arranged on purpose, the creatures, it is urged, having been suitably formed by the operation of chance, survived; otherwise they perished, and still perish, as Empedocles says of his "man-faced oxen".

Aristotle, however, dismissed the idea in, what seems to us, a rather supercilious manner:

> Such and suchlike are the arguments which may be urged in raising this problem; but it is impossible that this should really be the way of it. For all these phenomena and all natural things are either constant or normal, and this is contrary to the very meaning of luck or chance. No one assigns it to chance or to a remarkable coincidence if there is abundant rain in the winter, though he would if there were in the dog-days; and the other way about, if there were parching heat. Accordingly, if the only choice is to assign these occurrences either to coincidence or to purpose, and if in these cases chance coincidence is out of the question, then it must be purpose. But, as our opponents themselves would admit, these occurrences are all natural. There is purpose, then, in what is, and in what happens, in Nature.

By far the clearest statement of natural selection which has come down to us from classical antiquity is in *De rerum natura* of Lucretius. This famous passage, obviously based on Empedocles, describes how the unadapted monsters formed by chance perished, and that only those that were fit survived to procreate their like. Lucretius went even further and enumerated some of the characteristics which preserved certain animals in the struggle for existence and explained, by contrast, how domestic animals that were obviously unfitted for existence in the wild, had been preserved through the protection given them by man. From the translation of W. H. D. Rouse:

> Many were the monsters also that the earth then tried to make, springing up with wondrous appearance and frame: the hermaphrodite, between man and woman, yet neither, different from both; some without feet, others again bereft of hands; some found dumb also without a mouth, some blind without face; some bound fast with all their limbs

> adhering to their bodies, so that they could do nothing and go nowhere, could neither avoid mischief nor take what they might need. So with the rest of like monsters and portents that she made, it was all in vain: since nature denied them growth, and they could not attain the desired flower of age nor find food nor join by the ways of Venus. For we see that living beings need many things in conjunction, so that they may be able by procreation to forge out the chain of the generations: first there must be food, next there must be a way for the life-giving seeds to ooze through the frame and be discharged from the body, and that male and female be joined they must both have the means to exchange mutual pleasures.
>
> And many species of animals must have perished at that time, unable by procreation to forge out the chain of posterity: for whatever you see feeding on the breath of life, either cunning or courage or at least quickness must have guarded and kept that kind from its earliest existence; many again still exist, entrusted to our protection, which remain, commended to us because of their usefulness. Firstly, the fierce brood of lions, that savage tribe, has been protected by courage, the wolf by cunning, by swiftness the stag. But the intelligent dog, so light of sleep and so true of heart, beasts of burden of all kinds, woolly sheep also, and horned herds of oxen, all these are entrusted to men's protection, Memmius. For these have eagerly fled from the wild beasts, they have sought peace and the generous provision gained by no labour of theirs, which we give them as the reward of their usefulness. But those to which nature gives no such qualities, so that they could neither live by themselves at their own will, nor give us some usefulness for which we might suffer to feed them under our protection and be safe, these certainly lay at the mercy of others for prey and profit, being all hampered by their own fateful chains, until nature brought that race to destruction.

Thanks to Lucretius, the ability of nature to preserve certain types and eliminate others was clearly announced in the first century B.C. It would seem that a mature conception of natural selection might have developed as a logical consequence of such a promising start. Lucretius, however, was a mechanist; his philosophy tended to minimize the personal activity of the Gods, and, of course, it was received with hostility by the revealed religion of his day. The soon to emerge Christianity took over the perquisites of the older faiths and, becoming in turn the revealed religion, it continued to be hostile to Lucretius for, like the earlier religion, it looked upon him as a man who sought to limit the power of God. The abuse and calumny which Lucretius received for over fifteen hundred years is a matter of record, and because of the *odium theologicum*, his ideas were not given a fair hearing; not that a fair hearing would necessarily have advanced his mechanistic views, for an alternative concept fitted much better into the dominant philosophy of the period. Needless to say this concept was teleological.

It is not the purpose here to give even a brief history of teleology. However it would not be possible to explain the submergence of Lucretius' account of the origin of fitness in the animal kingdom without showing how the very adaptations he cited were used as evidences of design. The facts listed by Lucretius were, of course, well known; the conclusions drawn from them by others were very different. The growth and duration of the dominant teleological explanation can be shown by a few examples.

Maximus of Tyre was a philosopher who lived in the latter half of the second century A.D. Evidently, he had some conception of the universality of struggle and of the conquest of certain forms of life by others. In his dissertation on the question as to which were the more useful, soldiers or husbandmen, he listed a number of cases where nations of soldiers were free and brave but nations of husbandmen were slaves and cowards. He extended this concept down into the animal kingdom. From Dissertation XIII, translated by Taylor:[4]

> Proceed to animals: for here also you will behold liberty and slavery, and a life consisting of virtue and bonds. The ox ploughs, the horse contends: but if you change their works you will act illegally against nature. Timid animals feed on grass, bold animals hunt. The stag feeds on grass, the lion hunts. Jackdaws gather seeds, eagles hunt. But those animals that gather seed and live on grass are slaves, while those that hunt are free.

Maximus, however, was convinced that nature was orderly and planned and that anyone who endeavoured to do anything contrary to this planned economy was sinful and deserving of punishment. He recorded the instance of the Carthaginian youth who fed a lion cub on unlawful food and so changed his nature that he became a beast of burden. Maximus noted with approval that the Carthaginians put the impious youth to death for his illegal act. He then proceeded to list the characteristics of the animals which had been allotted to them for their preservation, having apparently had no conception of the possibilities that only those animals who were sufficiently endowed were preserved. From Dissertation XXXII:

---

[4] Taylor, Thomas, *The Dissertations of Maximus Tyrius*, London, 1804.

> As horses, therefore, for their safety are allotted the race, oxen labour, birds wings, lions strength, and other animals something else; in like manner a connate power which preserves the race is present with man. With respect to this power, it is necessary that it should be different from that of other animals; if, being man, he is to be saved, not by strength as lions, nor by the race as horse, and is not to carry burthens like an ass, nor to plough like an ox, nor to fly with birds, nor swim with fishes. But there is also a certain work peculiar to this animal, which preserves his life, if powers are distributed to animals according to the use of life, works according to powers, and instruments according to works, and the good they effect.

Again in Dissertation XIX he describes these characteristics as gifts.

> But with respect to other animals, one good is distributed to every herd for its preservation, and those of a similar species partake, with their like, of an equal life, and of one end, each with each, whether they fly, or walk, or creep, or live on flesh or grass; whether they are gregarious, tame, or wild; and whether they are horned, or without horns: and if you change the lives of these you act illegally towards nature.

Galen was definitely a teleologist and his great influence tended to strengthen the dominant philosophy of his day. He emphasized the use and function of the bodily parts in several of his works, particularly in *De usu partim*, *De formatione foetus* and *De placitis Hippocratis et Platonis*. In the latter work he definitely rejects the role of accident or chance in the production of any of the organs of our bodies.

The great Christian author, Lactantius (260–340 A.D.), proceeded directly to the refutation of Lucretius. He spared no words in his well known book, *On the Workmanship of God*, for he clearly deemed his opponent mistaken in a field where a mistake was a sin. From chapter 6:

> I cannot here be prevented from again showing the folly of Epicurus. For all the ravings of Lucretius belong to him, who, in order that he might show that animals are not produced by any contrivance of the divine mind, but, as he is wont to say, by chance, said that in the beginning of the world innumerable other animals of wonderful form and magnitude were produced; but that they were unable to be permanent, because either the power of taking food, or the method of uniting and generating had failed them. It is evident that, in order to make a place for his atoms flying about through boundless and empty space, he wished to exclude the divine providence. But when he saw that a wonderful system of providence is contained in all things which breathe, what

> vanity was it (O mischievous one!) to say that there had been animals of immense size, in which the system of production ceased.
>
> Since therefore, all things which we see are produced with reference to a plan—for nothing but a plan can effect this very condition of being born—it is manifest that nothing could have been born without a plan. For it was previously foreseen in the formation of everything, how it should use the service of the limbs for the necessaries of life; and how the offspring, being produced from the union of bodies, might preserve all living creatures by their several species.

Lactantius next describes the building of a house in a manner reminiscent of Paley's watch. He then proceeds:

> Why should any one suppose that, in the contrivance of animals, God did not foresee what things were necessary for living, before giving life itself? For it is manifest that life could not exist unless those things by which it exists were previously arranged.
>
> Therefore Epicurus saw in the bodies of animals the skill of a divine plan, but that he might carry into effect that which he had before imprudently assumed, he added another absurdity agreeing with the former. For he said that the eyes were not produced for seeing, nor the ears for hearing, nor the feet for walking; but that all the offices of these members arose from them after their production. I fear lest the refutation of such extravagant and ridiculous stories should appear to be no less foolish; but it pleases me to be foolish, since we are dealing with a foolish man, lest he should think himself too clever. What do you say, Epicurus? Were not the eyes produced for seeing? Why, then, do they see? Their use, he says, afterwards showed itself. Therefore they were produced for the sake of seeing, since they can do nothing else but see. Likewise, in the case of the other limbs, use itself shows for what purpose they were produced. For it is plain that this use could have no existence, unless all the limbs had been made with such arrangement and foresight that they might be able to have their use.
>
> For what, if you should say, that birds were not made to fly, nor wild beasts to rage, nor fishes to swim, nor men to be wise, when it is evident that living creatures are subject to that natural disposition and office to which each was created? But it is evident that he who has lost the main point itself of the truth must always be in error. For if all things are produced not by providence, but by a fortuitous meeting together of atoms, why does it never happen by chance, that those first principles meet together in such a way as to make an animal of such a kind, that it might rather hear with its nostrils, smell with its eyes, and see with its ears? For if the first principles leave no kind of position untried monstrous productions of this kind ought daily to have been brought forth, in which the arrangement of the limbs might be distorted and the use far different from that which prevails. But since all the limbs observe their own laws and arrangements, and the uses assigned to them, it is plain that nothing is made by chance, since a perpetual arrangement of the divine plan is preserved. But we will refute Epicurus at another time.

How completely Lactantius relied on the teleological explanation is shown by a passage in chapter 2. Nature may be "red in tooth and claw" but Providence has seen to it that there would be no lack of victims.

> Since He did not give that power of reason to the other animals, He provided beforehand in what manner their life might be more safe. For He clothed them all with their own natural hair, in order that they might more easily be able to endure the severity of frosts and cold. Moreover, He has appointed to every kind its own peculiar defense for the repelling of attacks from without; so that they may either oppose the stronger animals with natural weapons, or the feebler ones may withdraw themselves from danger by the swiftness of their flight, or those which require at once both strength and swiftness may protect themselves by craft, or guard themselves in hiding-places. And so others of them either poise themselves aloft with light plumage, or are supported by hoofs, or are furnished with horns; some have arms in their mouth—namely, their teeth—or hooked talons on their feet; and none of them is destitute of a defense for its own protection.
>
> But if any fall as a prey to the greater animals, that their race might not utterly perish, they have either been banished to that region where the greater ones cannot exist, or they have received a more abundant fruitfulness in production, that food might be supplied from them to the beasts which are nourished by blood, and yet their multitude might survive the slaughter inflicted upon them, so as to preserve the race.

St. Albertus Magnus (?1206–1280) was free from the violence of Lactantius and much more logical and subtle in his rebuttal of Empedocles. He described hybrid generation and admitted the possibility of the existence of the monsters of which Empedocles had written, but he looked upon them as accidents, as transgressions of nature aiming at a goal it was not good to reach. Nature herself was governed by design. The following quotation is from his *Physicorum* (Bk. II, tr. 3, ch. 3):

> A definite end, for which nature is working, is in both plants and animals. And from this consideration we shall contradict that of Empedocles, who said that all these things are chance occurrences both rightly and unrightly and monstrously [or unnaturally]: for if perchance, there be a monster among animals whose upper portion might be made of a man, and inferior portion might be made of the bull, which monster is a Minotaur; it is likewise necessary because such an accident might happen in plants according to Empedocles; thus, because at one time there may be generated a plant whose one part may be a vine and the other an olive: but however, this we do not see although it may be equally reasonable for this to happen in plants just as in animals according to Empedocles.

As we would expect, teleology dominated the natural theology of the seventeenth and eighteenth centuries. Even biologists who had considerable attainments to their credit interpreted the adaptation of animal and plant life as evidences of the design of a beneficent Providence. The best example of this attitude is to be found, perhaps, in the philosophy of the great pre-Linnean systematist, John Ray. The very title of one of Ray's books, *The Wisdom of God Manifested in the Works of Creation*, London, 1694, is sufficient to indicate how Ray explained the existence of fitness. From p. 135, Ed. of 1743:

> *Thirdly*, I shall remark the Care that is taken for the Preservation of the Weak, and such as are expos'd to the Injuries, and preventing the Increase of such as are noisome and hurtful; for as it is a Demonstration of the divine Power and Magnificence to create such Variety of Animals, not only great but small, not only strong and courageous, but also weak and timorous; so it is no less Argument of his Wisdom to give to these Means, and the Power and Skill of using them, to preserve themselves from the Violence of those. That of the Weak, some should dig Vaults and Holes in the Earth, as *Rabbets*, to secure themselves and their Young; others should be arm'd with hard Shells; others with Prickles; the rest that have no such Armature, should be endu'd with great Swiftness or Perspicacity; and not only so, but some also have their Eyes stand so prominent, as the *Hare*, that they can see as well behind as before them, that so they may have their Enemy always in their Eye; and long, hollow, moveable Ears, to receive and convey the least Sound, or that which comes from far, that they be not suddenly surprised or taken (as they say) napping. Moreover, it is remarkable, that in this Animal, and in the *Rabbet*, the Muscles of the Loins and the hind Legs are extraordinarily large in Proportion to the rest of the Body, or those of other Animals, as if made on purpose for Swiftness, that they may be able to escape the Teeth of so many Enemies as continually pursue and chase them; add hereto the Length of their hind legs, which is no small Advantage to them. . . .
>
> As for Sheep, which have no natural Weapons or Means to defend or secure themselves, neither Heels to run, nor Claws to dig, they are deliver'd into the Hand, and committed into the Care and Tuition of Man; and serving him for divers Uses, are nurished and protected by him; and so enjoying their Beings for a Time, by this Means propagate and continue their Species: So that there are none destitute of some Means to preserve themselves, and their Kind; and these Means so effectual, that notwithstanding all the Endeavours and Contrivances of Man and Beast to destroy them, there is not to this Day one *Species* lost of such as are mention'd in Histories and consequently and undoubtedly neither of such as were first created.
>
> Then for Birds of Prey, and rapacious Animals, it is remarkable what *Aristotle* observes, That they are all solitary, and go not in Flocks,

ΓΑΜΨΟΝΥΧΟΝ ΟΥΔΕΝ ΑΓΕΛΑΙΟΝ no Birds of Prey are gregarious. Again, that such creatures do not greatly multiply, ΤΕ ΓΑΜΨΟ ΙVΥΧΟΝ ΟΛΙΓΑΤΟΝ ΑΠΑΝΤΑ. They for the most part breeding and bringing forth but one or two, or at least, a few young ones at once: Whereas they that are feeble and timorous are generally multiparous; or, if they bring forth but a few at once, as *Pigeons*, they compensate that by their often breeding, *viz.* every Month but two throughout the Year; by this Means providing for the Continuation of their Kind. But for the Security of these rapacious Birds, it is worthy the noting, that because a Prey is not always ready, but perhaps they may fail of one some Days, Nature hath made them patient of long *inedia*, and besides when they light upon one, they gorge themselves so therewith, as to suffice for their Nourishment for a considerable Time.

Fourthly, I shall note the exact Fitness of the Parts of the bodies of Animals to every one's Nature and Manner of Living.

But one other example will be cited, the *Physico-Theology*, London, 1712, of William Derham, a book which had a considerable vogue in its day. Derham did not see in nature either population pressure or food shortage. From p. 181, Ed. of 1716:

3. Another wise Provision the Creator hath made relating to the Food of Animals is, that various Animals delight in various Food; some in Grass and Herbs; some in Grain and Seeds; some in Flesh; some in insects; some in this, some in that; some more delicate and nice, some voracious and catching at any thing. If all delighted in, or subsisted only with one sort of Food, there would not be sufficient for all; but every variety using various Food, and perhaps abhorring that of others, is a great and wise means that every Kind hath enough, and oftentimes somewhat to spare.

It deserves to be reckoned as an Act of the Divine Appointment, that what is wholsome Food to one, is nauseous, and as a Poyson to another; what is a sweet and delicate Smell and Taste to one, is foetid and loathsome to another. By which means all the Provisions the Globe affords are well disposed of. Not only every Creature is well provided for, but a due Consumption is made of those things that otherwise would encumber the World, lie in the way, corrupt, rot, stink, and annoy, instead of cherish and refresh. For our most useful Plants, Grain, and Fruits, would mould and rot; those Beasts, Fowls, and Fishes, which are reckoned among the greatest Dainties, would turn to Carrion, and Poyson us: Nay, those Animals, which are become Carrion, and many other things that are noysome, both on the Dryland, and in the Waters, would be more grievous Annoyances, and breed Diseases, was it not for the Provision which the infinite Orderer of the World hath made, by causing these things to be sweet, pleasant, and wholsome Food to some Creature or other, in the place where those things fall: to Dogs, Ravens, and other voracious Animals, for Instance, on the Earth; and to rapacious Fishes, and other Creatures inhabiting the Waters.

Thus is the World in some measure kept sweet and clean, and at the same time divers Species of Animals supplied with convenient Food.

Teleology furnished such a satisfactory explanation of fitness that natural selection was an unwanted hypothesis, a vagary, a trifle sinful perhaps, but not important enough to deserve more than a passing refutation. Indeed it was not until after Darwin published the *Origin of Species* that natural selection emerged as a respectable hypothesis and not until late in the nineteenth century did it succeed in chasing teleology from the biological field. Meanwhile only a few individuals had grasped its full implications, although a somewhat larger group had appreciated the importance of one or another of its basic postulates. The pre-Darwinian history of Natural selection must deal to a large extent with the contributions of this latter group.

We shall not be concerned here with the history of artificial selection; it was well understood by the Greeks and Romans who even extended its application to human beings. Roper[5] has covered this phase of the subject. The agricultural writings of the middle ages show that the importance of proper breeding stock was appreciated even if the agriculturists themselves did not always follow the best practices. In the early eighteenth century a number of plant breeders succeeded in producing many new and valuable varieties, while later in the century, Robert Bakewell started his experiments in animal breeding. Darwin was greatly interested in the practical results obtained by Bakewell and was undoubtedly influenced by him in developing the theory of natural selection. His recognition of the potency of artificial selection thus helped to establish natural selection as a factor in organic evolution. The subject, however, is too large to be included here.

When we endeavor to find the first contributors to the theory of natural selection after Lucretius we are faced with obvious gaps in the available records. We would expect little from the Latin World after the fall of Rome and little from the Byzantines. The Arabs may have contributed to the theory but we know practically nothing of what they did in the field. The writer is able to cite but a single pertinent excerpt from an Arabian scholar, a passage in the ninth century *Book of animals* of al-Jâhiz.

The following quotation (VI: 133) is from the Spanish translation of Miquel Asin Padacios (*Isis*, **14**: 20–54, 1930). Here al-Jâhiz describes the struggle for existence:

---

[5] Roper, A. G. *Ancient Eugenics*, Oxford, 1913.

> The rat goes out to look for its food, and is clever in getting it, since it eats all animals inferior to it in strength, such as little animals and small birds, the eggs and the young of the latter, and in general, the vermin which do not live in burrows or whose nests are flush with the earth. In its turn, the rat has to avoid snakes and birds and serpents of prey, who look for it in order to devour it. It must also be skillful in defending itself from the lizard and from the *herizo*, which are stronger than it is. The lizard is clever in hunting the snake and the fox. The latter in its turn, hunts all animals inferior to it. The mosquitoes go out to look for their food as they know instinctively that blood is the thing which makes them live. As soon as they see the elephant, hippopatamus or any other animal, they know that the skin has been fashioned to serve them as food; and falling on it, they pierce it with their probosces, certain that their thrusts are piercing deep enough and are capable of reaching down to draw the blood. Flies in their turn, although they feed on many and various things, principally hunt the mosquito which is the food which they like best. If it were not for the flies, the hum of mosquitoes during the day would be much greater. The star-lizard and the spider, called ant-lion, go out and hunt flies with the cleverest technique and the greatest dexterity. But in addition, flies disappear also through the medium of other causes, for example, they die upon eating in competition over the sweet morsels. All animals, in short, can not exist without food, neither can the hunting animal escape being hunted in his turn. Every weak animal devours those weaker than itself. Strong animals cannot escape being devoured by other animals stronger than they. And in this respect, men do not differ from animals, some with respect to others, although they do not arrive at the same extremes. In short, God has disposed some human beings as a cause of life for others, and likewise, he has disposed the latter as a cause of the death of the former.

In the seventeenth century some mention was made of several of the elements which make up natural selection. Francis Bacon (*New Atlantis*) told of the formation of new and useful varieties through hybridization, and he seemed to appreciate the necessity of selection in establishing the new forms. Thomas Hobbes described vividly the struggle for existence among human beings; in fact, he saw the establishment of the State as a necessary means of ameliorating the struggle. He would doubtless be completely at home in the world of 1940. The following quotations are from *Leviathan*, London, 1651, (Pt. I, ch. 13):

> Therefore if any two men desire the same thing, which nevertheless they cannot both enjoy, they become enemies; and in the way to their

> end, which is principally their own conservation, and sometimes their delectation only, endeavour to destroy, or subdue one another. And from hence it comes to pass, that where an invader hath no more fear, than another man's single power; if one plant, sow, build, or possess a convenient seat, others may probably be expected to come prepared with forces united, to dispossess, and deprive him, not only of the fruit of his labour, but also of his life, or liberty. And the invader again is in the like danger of another.
>
> And from this diffidence of one another, there is no way for any man to secure himself, so reasonable, as anticipation; that is, by force, or wiles, to master the persons of all men he can, so long, till he see no other power great enough to endanger him: and this is no more than his own conservation requireth, and is generally allowed. Also because there be some, that taking pleasure in contemplating their own power in the acts of conquest, which they pursue farther than their security requires; if others, that otherwise would be glad to be at ease within modest bounds, should not by invasion increase their power, they would not be able, long time, by standing only on their defence, to subsist. And by consequence, such augmentation of dominion over men being necessary to a man's conservation, it ought to be allowed him.

Hobbes next explains how men live in a state of war when they are not subjugated by a higher power (the State) which keeps them in awe, a war which is described as follows:

> To this war of every man, against every man, this also is consequent; that nothing can be unjust. The notions of right and wrong, justice and injustice have there no place. Where there is no common power, there is no law; where no law, no injustice. Force, and fraud, are in war the two cardinal virtues. Justice, and injustice are none of the faculties neither of the body, nor mind. If they were, they might be in a man that were alone in the world, as well as his senses, and passions. They are qualities, that relate to men in society, not in solitude. It is consequent also to the same condition, that there be no propriety, no dominion, no mine and thine distinct; but only that to be every man's, that he can get; and for so long, as he can keep it. And thus much for the ill condition, which man by mere nature is actually placed in; though with a possibility to come out of it, consisting partly in the passions, partly in his reason.

Hobbes' conclusions are realistic and to the point:

> And consequently it is a precept, or general rule of reason, that every man ought to endeavour peace, as far as he has hope of obtaining it; and when he cannot obtain it, that he may seek, and use, all helps, and advantages of war. The first branch of which rule, containeth the first, and fundamental law of nature; which is, to seek peace, and follow it. The second, the sum of the right of nature; which is, by all means we can, to defend ourselves.

Sir Matthew Hale would probably have secured a prominent position in the literature of evolution if all of his ideas which were relevant to the subject had been recorded in a single passage. As it is we must search through his work, *The Primative Origination of Mankind*, London, 1677, to collect a few sentences here and there which bear on the subject. Hale believed in degeneration (mutation), the inheritance of acquired characters and the formation of new varieties by hybridization. The following quotation shows that he also recognized the struggle for existence and that he anticipated Malthus in listing the checks to an excessive increase of population. From p. 211–212:

> So among Brutes, Birds, Fishes and Insects there is a continual invading and prevalence of the more powerful, active and lively, over the more weak, flegmatick, and unactive Creatures; the Bear, Lion, Wolf, Dog, Fox, *etc.* pursue the Sheep, Oxen, Hare, Coney, *etc.* and prey upon them: the like is evident among Birds and Fishes, and generally Insects, being the weaker and more unconsiderable parts of Nature.

Hale ascribed the population checks to the skillful planning of a wise Providence, which succeeds in preserving the balance of nature. He proceeds:

> And upon this seemingly impertinent Diversion touching the Reductions and Correctives of these inferior Animals, there may seem to be collected reasonably an analogical Inference of the like means of the Correctives of the Generations of Mankind; and that although in an ordinary course of Human Productions the Increase surmounts the Decay, yet there may be reasonably supposed such Periodical Corrections as might fairly keep the state of Mankind in a mediocrity and equability although it should be supposed the Generations of Mankind had been Eternal.
>
> And although these Correctives may not happen every Day, or every Year in the ordinary course of things, and therefore may be called extraordinary, because they are less ordinary than the common Causalties of Mankind, as Sickness or Accident that happens to this or that individual Person promiscuously, yet they are in truth no more extraordinary, than a cold Winter is extraordinary; which although it is not every Day, nor doth it happen every Year possibly in an equal Degree, yet it is no extraordinary thing in Nature, if it happens once in 5, or 10, or 20 years.
>
> Having therefore considered these Correctives in the inferior Animal Nature, I shall now search out what may be those Correctives, that may be applicable to the Reductions of the Generations of Mankind to an Equability, or at least to keep it within such bounds as may keep it from surcharging the World; whereby if in the Period of 2, or 3, or 4000 years it may grow too luxuriant, yet it may in probability be so far abated, as

> may allow it an Increase of the like number of Years to attain its former proportion. So that by these Prunings there may be a consistency of the Numbers of Mankind, with an eternal succession of Individuals.
>
> Those Reductions that may be supposed effectual for these ends, and such as the course of Mankind seem to have great Experience of are, 1. Plagues and Epidemical Diseases: 2. Famines: 3. Wars and Internecions: 4. Floods and Inundations: 5. Conflagrations.

Hale offers a compromise between special creation and the continuous formation of new species (p. 304):

> 2. That all the *Species* of perfect Animals of all kinds were constituted in their several Sexes in the fifth and sixth day of the Creation, but yet we must not think that all those kinds which we now see were at first created, but only those primitive and radical *Species*: How many sorts of Animals do we now see, that yet possibly are not of the same *Species*, but have accidental diversifications, as we may observe in the several Shapes and Bodies of Dogs, Sheep, Pyes, Parots? which possibly at first were not so diversified; some variation of the same *Species* happen by mixt Coition, some by diversity of Climates, and other accidents.

During the eighteenth century scholars continued to recognize the separate postulates which together make up our modern conception of natural selection, and the many incidental references to one or another of its elementary propositions show that it would be only a matter of time for the whole theory to emerge. The effects of sexual selection on human beings was noted at this time by Jean Chardin and described by him in *Voyages in Perse*, London, 1711. From Vol. IV, p. 98, Ed. of 1723:

> That the blood of the Persians is naturally gross, appears from the Guebres, who are a remnant of the ancient Persians, and are an ugly, ill-made, rough-skinned people. This is also apparent from the inhabitants of the provinces in the neighbourhood of India, who never formed alliances with any other tribes. But, in the other parts of the kingdom, the Persian blood is now highly refined by frequent intermixtures with the Georgians and Circassians, two nations which surpass all the world in personal beauty. There is hardly a man of rank in Persia who is not born of a Georgian or Circassian mother; and even the king himself is commonly sprung, on the female side from one or other of these countries. As it is long since this mixture commenced, the Persian women have become very handsome and beautiful, though they do not rival the ladies of Georgia. The men are generally tall and erect, their complexion is ruddy and vigorous, and they have a graceful air and an engaging deportment. The mildness of the climate, joined to their temperance in living, has a great influence in improving their personal beauty. This quality they inherit not from their ancestors; for, without

the mixture above mentioned, the men of rank in Persia, who are descendants of the Tatars, would be extremely ugly and deformed.

Buffon's contribution to the theory of evolution is well known, for he was undoubtedly the most famous evolutionist of his day. The passages from his works to be cited here show that he understood the implications of population pressure, thus anticipating Malthus, and that he had a very clear appreciation of the struggle for existence.

> Nature, in general, seems to have a greater tendency to life than death, and to organize bodies as much as possible; the multiplication of germs, which may be infinitely increased, is a proof of it; and we may assert with safety, that if all matter is not organized, it is because organized beings destroy each other; or we can augment as much as we please the quantity of living and vegetating beings, but we cannot augment the quantity of stones or other inanimate matters. This seems to indicate that the most common work of Nature is the production of the organic part, and in which her power knows no bounds.
>
> To render this intelligible, let us make a calculation of what a single germ might produce. The seed of an elm, which does not weigh the hundredth part of an ounce, at the end of 100 years will produce a tree whose volume will be 60 cubic feet. At the tenth year this tree will have produced 1000 seeds, which being all sown, at the end of 100 years would each have also a volume equal to 60 cubic feet. Thus in 110 years there is produced more than 60,000 cubic feet of organized matter; 10 years more there will be 10,000,000 of fathoms, without including the 10,000 increased every year; and ten years after there will be three times that number; thus in 130 years a single shoot will produce a volume of organized matter, which would fill up a space of 1000 cubic leagues; 10 years after it could comprehend a 1,000,000, and in 10 years more 1,000,000 times 1,000,000 cubic leagues; so that in 150 years the whole terrestrial globe might be entirely converted into one single kind of organized matter. In this production of organized body Nature would know no bounds, if it were not for the resistance of matters which are not susceptible of organization, and this proves that she does not incline to form inanimate but organized beings, and that in this she never stops but when irresistible inconveniences are opposed thereto. What we have already said on the seed of an elm may be said of any other; and it would be easy to demonstrate, that if we were to hatch every egg produced by hens for the space of 30 years, there would be a sufficient number of fowls to cover the whole surface of the earth.
>
> In fertile ages, and when population is the greatest, the whole surface of the earth seems to be covered with men, domestic animals and useful plants. But in the times of famine and depopulation, the ferocious animals, poisonous insects, parasitical plants, and useless herbs resume, in their turn, dominion over the earth. To man these changes are material, but to Nature they are perfectly indifferent. The silk worm so

> inestimable to the former, is to the latter only a caterpillar of the mulberry tree. Though this caterpillar, which so materially assists in the supply of our luxuries, should disappear; though the plants, from which our domestic animals procure their nourishment, should be devoured by other caterpillars; though still others should destroy the substance of our corn before the harvest; in short, though men and the larger animals should be starved by the inferior tribes, Nature would not be less abundant nor less alive; she never protects one at the expense of another, but especially supports the whole.
>
> Let us consider any of the inferior species which serve as food to others; herrings, for example, present themselves in millions to our fishermen, and after having fed all the monsters of the northern sea, they contribute to the subsistence of all the nations in Europe for a certain part of the year. If prodigious numbers of them were not destroyed, what would be the effects of their prodigious multiplication? By them alone would the whole surface of the sea be covered. But their numbers would soon prove a nuisance; they would corrupt and destroy each other. For want of sufficient nourishment their fecundity would diminish; by contagion and famine they would be equally destroyed; the number of their own species would not be increased, but the number of those that feed upon them would be diminished. As this remark is alike applicable to any other species, so it is necessary they should prey upon each other; the killing of animals, therefore, is both a lawful and innocent custom, since it is founded in nature, and it is upon that seemingly hard condition they are brought into existence. . . .
>
> It is by the right of conquest, however, that he [man] reigns; he rather enjoys than possesses, and it is by perpetual activity and vigilance that he preserves his advantage; if those are neglected everything languishes, changes, and returns to the absolute dominion of Nature, she resumes her power, and destroys the operations of man; envelopes with moss and dust his most pompous monuments, and in the progress of time entirely effaces them, leaving him to regret having lost by his negligence what his ancestors have acquired by their industry.

One of Buffon's famous contemporaries was Denis Diderot, whose evolutionary teachings biologists have overlooked for some unknown reason. His statement of the case for evolution is so clear and accurate that it almost seems that we would be forced to accept his conclusions as a logical necessity even in the absence of the evidence collected since his time. There is space here only for a passage which very definitely urges natural selection as an alternate to the teleological explanation of adaptation. The quotation is in the form of a dialogue between Mr. Holmes and Mr. Saunderson, a blind man, and occurs in *Lettre sur les aveugles*, published in 1749. From Diderot's *Oeuvres*, Paris, 1818, vol. I, p. 319: Mr. Saunderson is speaking:

"Therefore imagine, if you please, that the order which is so striking has always existed; but let me believe that there is none, and that if we go back to the birth of things and time, and if we could sense matter moving and chaos unravelling, we should meet a multitude of unformed beings for each several well-organized beings. If I have nothing to object to you of on the condition of present things, I can at least question you as to their past condition. I can ask you, for example, who told you at Leibnitz, at Clarke, and at Newton that in the first place in the formation of animals some were not without heads and others without feet. I can maintain to you that this one had no stomach and that one no intestines; that such in whom a stomach, palate, and teeth seemed to permit continuation were exterminated by some defect of heart or lungs; that monsters appeared, and that there only remained those whose mechanisms implied no important contradiction, and which could subsist by themselves and perpetuate themselves.

"This supposed, if the first man had had a closed larynx, had lacked the suitable food, had failed through the parts of generation, had not met his mate, or had spread to another species, M. Holmes, what would have become of humanity? He would have been enveloped in the general purification of the universe; and this proud being who calls himself man, dissolved and dispersed between the molecules of matter, would have remained perhaps forever, in the ranks of possibilities.

"If there had never been malformed creatures, you would not fail to insist that there will never be any; and that I am throwing myself into a chimerical hypothesis: but order is not so perfect," continued Saunderson, "that there do not still appear, from time to time, monstrous productions."

No history of the ideas which led to our present evaluation of natural selection would be complete without some mention of a little known essay by Benjamin Franklin. Franklin is really the source of Darwin's inspiration, for he gave Malthus the clue to the theory of population we now call Malthusian, and Malthus, as has been stated, gave Darwin the clue which led to the discovery of natural selection. The essay in question is entitled *Observations concerning the Increase of Mankind and the Peopling of Countries.* It was written in 1751 and published in Boston in 1755. Its contents and ideology will probably come as a surprise to many, for it expresses an attitude which we do not associate with Franklin. Sections 22, 23 and 24 are relevant enough to be quoted in full.

22. There is, in short, no bound to the prolific nature of plants or animals, but what is made by their crowding and interfering with each other's means of subsistence. Was the face of the earth vacant of other plants, it might be gradually sowed and overspread with one kind only, as, for instance, with fennel; and, were it empty of other inhabitants, it

might in a few ages be replenished from one nation only, as, for instance, with Englishmen. Thus, there are supposed to be now upwards of one million English souls in North America (though it is thought scarce eighty thousand has been brought over sea), and yet perhaps there is not one the fewer in Britain, but rather many more, on account of the employment the colonies afford to manufacturers at home. The million doubling, suppose but once in twenty-five years will, in another century, be more than the people of England, and the greater number of Englishmen will be on this side of the water. What an accession of power to the British empire by sea as well as land! What increase of trade and navigation! What numbers of ships and seamen! We have been here but little more than one hundred years, and yet the force of our privateers in the late war, united, was greater, both in men and guns, than that of the whole British navy in Queen Elizabeth's time. How important an affair then to Britain is the present treaty for settling the bounds between her colonies and the French, and how careful should she be to secure room enough, since on the room depends so much the increase of her people.

23. In fine, a nation well regulated is like a polypus; take away a limb, its place is soon supplied; cut it in two, and each deficient part shall speedily grow out of the part remaining. Thus, if you have room and subsistence enough, as you may, by dividing, make ten polypuses out of one, you may of one make ten nations, equally populous and powerful; or rather increase a nation ten fold in numbers and strength.

The following paragraph is of especial interest in connection with the present (1940) German estimate of Germans.

And since detachments of English from Britain, sent to America, will have their places at home soon supplied and increase so largely here; why should the Palatine boors be suffered to swarm into our settlements, and, by herding together, establish their language and manners, to the exclusion of ours? Why should Pennsylvania, founded by the English, become a colony of aliens, who will shortly be so numerous as to Germanize us instead of our Anglifying them, and will never adopt our language or customs any more than they can acquire our complexion?

24. Which leads me to add one remark, that the number of purely white people in the world is proportionably very small. All Africa is black or tawny; Asia chiefly tawny; America (exclusive of the newcomers) wholly so. And in Europe, the Spaniards, Italians, French, Russians, and Swedes are generally of what we call a swarthy complexion; as are the Germans also, the Saxons only excepted, who, with the English, make the principal body of white people on the face of the earth. I could wish their numbers were increased. And while we are, as I may call it, scouring our planet, by clearing America of woods, and so making this side of our globe reflect a brighter light to the eyes of inhabitants in Mars or Venus, why should we, in the sight of superior beings, darken its people? Why increase the sons of Africa by planting them in America,

> where we have so fair an opportunity, by excluding all blacks and tawnys, of increasing the lovely white and red? But perhaps I am partial to the complexion of my country, for such partiality is natural to mankind.

Jean-Jacques Rousseau, in his ecstatic dream of primitive man, actually saw natural selection as an agent for improving the human species. The following citation is from *Discours sur l'origine et les fondamens de l'inégalitè parmi les hommes*, Amsterdam, 1755. The italics are the writer's:

> Accustomed from their infancy to the inclemencies of the weather, and the rigour of the seasons; inured to fatigue, and forced naked and unarmed, to defend themselves and their prey from other ferocious animals, or to escape them by flight, mankind would acquire a robust and almost unalterable temperament of body. In the mean while, the children, bringing with them into the world the excellent constitution of their parents, and then confirming it by the same exercises which first produced it, would thus acquire all that strength and vigour, of which the human frame is capable. *Nature in this case treats them exactly as Sparta treated the children of her citizens: those of them who came well formed into the world, she renders strong and robust, and destroys all the rest;* differing totally in this respect from our modern communities; in which the state, by permitting children to become burthensome to their parents, murders them all without distinction even in the mother's womb.

Lovejoy[6] has called attention to the fact that P. L. M. de Maupertius saw in natural selection an explanation of adaptation. Maupertius did not consider natural selection the only possible explanation but he showed that the biological sciences were not compelled to resort to teleology to account for the existence of fitness.

> . . . May we not say that, in the fortuitous combination of the productions of Nature, since only those creatures *could* survive in whose organization a certain degree of adaptation was present, there is nothing extraordinary in the fact that such adaptation is actually found in all those species which now exist? Chance, one might say, turned out a vast number of individuals; a small proportion of these were organized in such a manner that the animals' organs could satisfy their needs. A much greater number showed neither adaptation nor order; these last have all perished. . . . Thus the species which we see today are but a small part of all those that a blind destiny produced.

The position of Immanuel Kant as a forerunner of Darwin has given rise almost to a literature of its own. As early as 1875,

---

[6] Lovejoy, Arthur. "Some Eighteenth Century Evolutionists." *Pop. Sci. Monthly*, **65**: 238–251, 323–340 (1904).

Fritz Schultz published *Kant und Darwin*, and Ernst Haeckel stated ". . . and it is again our great Königsberg philosopher, Immanuel Kant, in whom we find the first ideas of this theory already a century before Darwin" (*History of Creation*, ed. of 1896, p. 172). Kant, it's true, seems to have played all around the idea of natural selection but the idea always just eluded him. In his *Physical Geography* (1757) he showed that he realized the importance of both artificial selection and heredity, and in his treatise *On the Different Races of Man* (1775), he held that it was the concept of a careful selection of special births which made possible the notion that an improved race of men could be bred. He visualized the struggle for existence for he stated in his *Pragmatisch Anthropologie:*

> Nature has placed the germ of dissension in the human race, and this becomes the means by which the amelioration of the race is accomplished by progressive culture. The inner and outward struggle is the impetus wherewith man passes from a rude state of nature into that of a citizen, just as in the case of a piece of machinery, where two opposite forces thwart each other by friction, but are nevertheless kept in motion by the blow or pull of other forces.

It is not clear whether Kant visualized the improvement wrought by the struggle for existence as due to a liquidation of the unfit or to the fact that it kept the individuals involved well exercised and in good athletic trim. The balance of evidence indicates that Kant never grasped the full logical implications of natural selection, however, and he even expressed the opinion that the problem which natural selection solves could not be solved. Osborn [7] has called attention to the fact that Kant's views were rather mechanistic in his earlier years but became more teleological later. He could not conceive of the suitability of form in the animal and plant kingdom without purpose. In the following passage, quoted from the English translation in Osborn's *From the Greeks to Darwin*, Kant simply gives up:

> It is quite certain that we cannot become sufficiently acquainted with organized creatures and their hidden potentialities by aid of purely mechanical natural principles, much less can we explain them; and this is so certain, that we may boldly assert that it is absurd for man even to conceive such an idea, or to hope that a Newton may one day arise even to make the production of a blade of grass comprehensible, according to

---

[7] Osborn, Henry Fairfield. *From the Greeks to Darwin*, New York, 1896.

natural laws ordained by no intention; such an insight we must absolutely deny to man.

Charles Bonnet recognized population pressure in the crowding together of individuals and species in nature (*Contemplation de la nature*, Amsterdam, 1762). He seems to have looked upon the arrangement as the working out of a rather clever plan. On the other hand the excentric James Burnet (Lord Monboddo) saw a definite stimulus to improvement induced by the struggle for existence. We can list Burnet with those who anticipated Malthus on the strength of a passage in his great work, *Of the Origin and Progress of Language*, Edinburgh, 1773–92. From Vol. I, p. 39:

> But the most fruitful country may be overstocked with any animal and particularly with man, who I believe is maintained with more difficulty, even in his natural state than other animals of much larger size: For I hold, that he cannot subsist upon herbage or foliage alone, but must have seeds, fruits, roots, or flesh. And it is to be considered, that man must have multiplied very much in his natural state, as he likewise does in the first stages of society. Now, when men were so multiplied that the natural fruits of the earth could not maintain them, they were under a necessity to practice one or other of the following methods; either to disperse, and go in search of other countries, where they might subsist more at their ease. But this in many cases might be impracticable: For the countries around them might be, and in the process of time certainly would be, as much overstocked as theirs; or they might be hindered by seas, great rivers, or impassable desarts. To all which may be added, the natural aversion that every animal has to quit its native country, and the haunts to which it has been accustomed. Or, 2nd, They might prey upon other animals, or upon one another. But this, besides the danger of it, would hardly be practicable by man solitary, unassisted by arts, and without other weapons than those which nature has given him. Or, *lastly*, They must associate and provide in common what singly they could not procure. And this last method, it is natural to think, so sagacious an animal as man would prefer to either of the other two.

We would naturally look into the anti-teleological writings of David Hume for some use of the theory of natural selection to explain adaptation. Hume did so use the theory but his conception of it was not as precise nor his statement of the case as clear as that of his predecessors, Diderot and Maupertius. The passage in question occurs in his posthumous work, *Dialogues concerning Natural Religion*, published in 1779, three years after his death. From p. 490 (Ed. of London, 1788):

> It is in vain therefore, to insist upon uses of the parts in animals or vegetables, and their curious adjustments to each other. I would fain know how an animal could subsist, unless its parts were so adjusted? Do we not find, that it immediately perishes wherever this adjustment ceases, and that its matter corrupting tries some new form? It happens, indeed, that the parts of the world are so well adjusted, that some regular form immediately lays claim to this corrupted matter: and if it were not so, could the world subsist? Must it not dissolve as well as the animal, and pass through new positions and situations: till in a great, but finite succession, it falls at last into the present or some such order?

Hume thus saw that non-adjusted animals could not exist to confront the teleogists and that an unbalanced world itself would not be in any state to support teleologists who could speculate about its unbalance. He also saw that there was a struggle for existence. From p. 500:

> The whole earth, believe me, Philo, is cursed and polluted. A perpetual war is kindled among all living creatures. Necessity, hunger, want, stimulated the strong and courageous: Fear, anxiety, terror, adjitates the weak and infirm.

Lovejoy, in the article cited, also called attention to the contribution of Herder to the conception of population pressure and of the struggle for existence. Herder clearly saw that population pressure made for war in the animal kingdom and that man was able to emerge triumphant over the other animals because he was eminently successful in this struggle. Although he personified Nature he saw clearly the struggle itself might cause the extinction of some species if a changed environment rendered their past adaptations unfit. Whatever the outcome, however, it was obviously in harmony with Nature's "intention". Herder's anticipation of Malthus was included in his *Ideen zur Geschichte der Menscheit*, 1784. The quotations are from the English Edition, *Outlines of a Philosophy of the History of Man*, London, 1803. From Vol. I, p. 53:

> Nature employs germes, she employs an infinite number of germes, because in her grand progress she promotes a thousand ends at once. She must also calculate upon some loss; as every thing is crowded, and nothing finds room completely to develope itself. But that, amid this apparent prodigality, the essential, and the first, fresh powers of life, with which she must necessarily prevent all accidents in the course of being so thronged, might never fail, she made the season of youth the season of love, and kindled her torch with the most subtile and active fire between earth and Heaven.

From p. 61–63:

. . . One species he [Man] must tame: with another he must long contend. Some escape his dominion: others wage with him eternal war. In short, every species extends it's possission of the Earth in proportion to it's capacity, cunning, strength, or courage.

It is not here the question, whether man have reason, and beasts have none. If they have not, they have some other advantages: for assuredly Nature has left none of her offspring unprotected. Were a creature neglected by her, from whom could it obtain succor? Since the whole creation is at war, and the most opposite powers are found so close to each other. Here godlike man is annoyed by snakes, there by vermin, here a shark devours him, there a tiger. Each strives with each, as each is pressed upon; each must provide for his own subsistence, and defend his own life.

Why acts Nature thus and why does she thus crowd her creatures one upon another? Because she would produce the greatest number and variety of living beings in the least space, so that one crushes another, and an equilibrium of powers can alone produce peace in the creation. Every species cares for itself, as if it were the only one in existence: but by it's side stands another, which confines it within due bounds: and in this adjustment of opposing species creative Nature found the only means of maintaining the whole. She weighed the powers, she numbered the limbs, she determined the instincts of the species toward each other; and left the earth to produce what it was capable of producing.

I concern myself not, therefore, whether whole species of animals have perished from the face of the Earth. Has the mammoth disappeared? So have giants. When these existed, the relations between the several creatures were different: . . . .

. . . Thus in the present constitution of our Earth no species of animals has been lost: though I question not but others may have existed, when it's constitution was different; and if at any future period Art or Nature should completely change it, a different relation between living creatures would take place.

Man, in short, entered an inhabited world. All the elements, rivers, and morasses, earth and air, were filled or filling, with living creatures: and he had to make room for his dominion by his godlike qualities, skill and power.

Although the theory of sexual selection is not an integral part of natural selection it was used by Darwin as a most important supplement to his main thesis and developed in great detail in his *Descent of Man.* Three early uses of this theory are included as samples of the status of the theory before Darwin brought it to the attention of biologists in general. The first of these passages has been quoted (Chardin). The second is taken from *An Essay on the Causes of Varieties of Complexion and Figure in the Human*

Species, Philadelphia, 1787 by Samuel Stanhope Smith, President of the College of New Jersey. From the edition of 1810, p. 188:

> . . . The superior ranks, with few exceptions, will generally excel, in the beauty of their form and complexion, not only because they enjoy in a higher degree, other advantages which have been already pointed out as contributing to this end, but because they have it more in their power to form connexions in marriage among the most beautiful of the sex. The Persian nobility, who are of Tartarian origin, have, in consequence of their removal into a more favorable climate, and their having adopted the manners of a civilized people, acquired juster ideas of the perfection of the human form than they possessed in their primitive seats. Hence, being led to seek the most beautiful women in marriage, they have exchanged the harsh features, and disproportioned figures of the Tartar ancestors, for a stature tall, and elegant, and a form and expression of countenance noble and commanding. The Turkish families of fortune have, in like manner, improved the physical character of their race. And if we may ascribe any truth to the portraits drawn by Roman historians of the ancestors of the present nations of Europe, we must acknowledge that the refinement of manners, and the improvements in the state of society, which have been introduced in modern times among their descendents, have contributed also to produce a proportional improvement in their features, and their persons.

William Smellie who wrote *Philosophy of Natural History*, Edinburgh, 1790–1799, described both population pressure and the struggle for existence. A judicious selection from his work would make him appear to be a real precursor of Darwin (pp. 391–396):

> The hostilities of animals, mankind not excepted, give rise to mutual improvement. Animals improve, and discover a superiority of parts, in proportion to the number of enemies they have to attack or evade. . . . Hostilities, in some instances, seem to arise, not from a natural antipathy of one species to another, but from a scarcity of food. . . . A profusion of animal life seems to be the general intention of Nature.

These facts however were interpreted to show that nature was planned, beneficent, and wise. Thus:

> But the greatest possible extension of life would still be wanting, if animals did not prey upon each other. If all animals were to live upon vegetables alone, many species, and millions of individuals, which now enjoy life and happiness, could have no existence; for the productions of the earth would not be sufficient to support them. But, by making animals feed upon each other, the system of animation and of happiness is extended to the greatest possible degree. In this view, Nature, instead of being cruel and oppressive, is highly generous and beneficent.

With Thomas Robert Malthus, we reach the acknowledged inspirer of both Charles Darwin and Alfred Russel Wallace. Though Malthus has always stood high in the estimation of biologists he has usually been represented as the devil's advocate in sociological literature. The numerous refutations of his doctrines some day should furnish interesting material for the history of science. It may not be out of place to give here a brief estimate of the validity of his conclusions, for most criticisms of Malthus are still based on political rather than scientific considerations. This may be done most simply by a series of numbered propositions. First, the number of individuals in every species of animal and plant does tend to increase, whenever it is not checked, at a rate represented by a geometrical progression. Second, in a state of nature the available food supply tends to remain constant if we disregard, of course, minor fluctuations. As a consequence, all populations in nature remain relatively constant, *i.e.*, they change very slowly and over a great period of time. (But we must except from our calculations those occasions where some force has temporarily upset the balance of nature.) Third, the food supply for human beings has shown a considerable increase in the immediate past and probably will show a further increase for some time to come. Fourth, the Malthusian "law," population pressure, does restrain the natural increase of all species of animals and plants with the following possible exceptions: (1) plants in arid and sub-arid regions where the vegetation is sparse and the competition between individual plants is reduced to a minimum, (2) that fraction of the human race which practices birth control.

Two short excerpts from the first chapter of the *Essay* will show how pertinent Malthus' work was to Darwin's conception of natural selection.

> The cause to which I allude is the constant tendency in all animated life to increase beyond the nourishment prepared for it.
>
> It is observed by Dr. Franklin, that there is no bound to the prolific nature of plants or animals but what is made by their crowding and interfering with each other's means of subsistence. Were the face of the earth, he says, vacant of other plants, it might be gradually sowed and overspread with one kind only, as, for instance, with fennel; and were it empty of other inhabitants, it might in a few ages be replenished from one nation only, as, for instance, with Englishmen.
>
> This is incontrovertibly true. Throughout the animal and vegetable kingdoms Nature has scattered the seeds of life abroad with the most

profuse and liberal hand; but has been comparatively sparing in the room and the nourishment necessary to rear them. The germs of existence contained in this earth, if they could freely develop themselves, would fill millions of worlds in the course of a few thousand years. Necessity, that imperious, all-pervading law of nature, restrains them within the prescribed bounds. The race of plants and the race of animals shrink under this great restrictive law; and man cannot by any efforts of reason escape from it.

In plants and irrational animals, the view of the subject is simple. They are all impelled by a powerful instinct to the increase of their species, and this instinct is interrupted by no doubts about providing for their offspring. Whenever, therefore, there is liberty, the power of increase is exerted, and the superabundant effects are repressed afterwards by want of room and nourishment.

The effects of this check on man are more complicated. Impelled to the increase of his species by an equally powerful instinct, reason interrupts his career, and asks him whether he may not bring beings into the world for whom he cannot provide the means of support. If he attends to this natural suggestion, the restriction too frequently produces vice. If he hears it not, the human race will be constantly endeavoring to increase beyond the means of subsistence. But as, by that law of our nature which makes food necessary to the life of man, population can never actually increase beyond the lowest nourishment capable of supporting it, a strong check on population, from the difficulty of acquiring food, must be constantly in operation. The difficulty must fall somewhere, and must necessarily be felt in some or other of the various forms of misery, or the fear of misery, by a large portion of mankind.

That population has this constant tendency to increase beyond the means of subsistence, and that it is kept to its necessary level by these causes, will sufficiently appear from a review of the different states of society in which man has existed. But, before we proceed to this review, the subject will perhaps be seen in a clearer light, if we endeavour to ascertain what would be the natural increase of population, if left to exert itself with perfect freedom; and what might be expected to be the rate of increase of productions of the earth, under the most favorable circumstances of human industry.

It will be allowed that no country has hitherto been known where the manners were so pure and simple, and the means of subsistence so abundant, that no check whatever has existed to early marriages from the difficulty of providing for a family, and that no waste of the human species has been occasioned by vicious customs, by towns, by unhealthy occupations, or too severe labour. Consequently in no state that we have yet known, has the power of population been left to exert itself with perfect freedom.

Whether the law of marriage be instituted or not, the dictate of nature and virtue seem to be an early attachment to one woman; and where there were no impediments of any kind in the way of a union to which such an attachment would lead, and no causes of depopulation

> afterwards, the increase of the human species would be evidently much greater than any increase which has been hitherto known.
>
> In the northern states of America, where means of subsistence have been more ample, the manners of the people more pure, and the checks to early marriages fewer, than in any of the modern states of Europe, the population has been found to double itself, for above a century and a half successively, in less than twenty-five years. Yet, even during these periods, in some of the towns, the deaths exceeded the births, a circumstance which clearly proves that, in those parts of the country which supplied the deficiency, the increase must have been much more rapid than the general average.
>
> . . . .
>
> In this supposition no limits whatever are placed to the produce of the earth. It may increase forever, and be greater than any assignable quantity; yet still the power of population being in every period so much superior, the increase of the human species can only be kept down to the level of the means of subsistence by the constant operation of the strong law of necessity, acting as a check upon the greater power.

Malthus was acquainted with the effectiveness of artificial selection in the improvement of domestic animals of eighteenth century England. This knowledge needed only to be combined with his own conception of population pressure to give him the essential elements of natural selection and even of organic evolution. But Malthus never made the synthesis. It seems a touch ironic but he was obviously prevented from anticipating Darwin by his own righteous indignation. His famous *Essay* was written to refute the wishful thinking of Godwin and to destroy the sweetness and light of Condorcet. Condorcet was all for perfecting the human race but Malthus saw the difficulties of this particular program. In emphasizing the limitations to the improvement which artificial selection could make in domestic stock, Malthus missed the idea of evolution. From Book III: chapter 1:

> I have been told that it is a maxim among some of the improvers of cattle that you may breed to any degree of nicety that you please, and they found this maxim upon another, which is that some of the offspring will possess the desirable qualities of the parents in a greater degree. In the famous Leicestershire breed of sheep, the object is to procure them with small heads and small legs. Proceeding upon these breeding maxims, it is evident that we might go on till the heads and legs were evanescent quantities; but this is so palpable an absurdity, that we may be quite sure the premises are not just, and that there really is a limit, though we cannot see it or say exactly where it is. In this case the point of the greatest degree of improvement, or the smallest size of the

head and legs, may be said to be undefined; but this is very different from unlimited or from indefinite, in M. Condorcet's acceptation of the term. Though I may not be able in the present instance to mark the limit at which further improvement will stop, I can very easily mention a point at which it will not arrive. I should not scruple to assert that were the breeding to continue forever, the head and legs of these sheep would never be so small as the head and legs of a rat.

It cannot be true therefore that among animals some of the offspring will possess the desirable qualities of the parents in a greater degree, or that animals are indefinitely perfectible.

The progress of a wild plant to a beautiful garden flower is perhaps more marked and striking than anything that takes place among animals, yet even here it would be the height of absurdity to assert that the progress was unlimited or indefinite. One of the most obvious features of the improvement is the increase of size. The flower has gradually grown larger by cultivation. If the progress were really unlimited, it might increase *ad infinitum;* but this is so gross an absurdity that we may be quite sure that among plants as well as among animals there is a limit to improvement, though we do not know exactly where it is. It is probable that the gardeners who contend for flower-prizes have often applied strong dressing without success. At the same time it would be highly presumptuous in any man to say that he had seen the finest carnation or anemone that could ever be made to grow. He might however assert, without the smallest chance of being contradicted by a future fact, that no carnation or anemone could ever by cultivation be increased to the size of a large cabbage; and yet there are assignable quantities greater than a cabbage. No man can say that he has seen the largest ear of wheat, or the largest oak, that could ever grow; but he might easily, and with perfect certainty, name a point of magnitude at which they would not arrive. In all these cases, therefore, a careful distinction should be made between an unlimited progress and a progress where the limit is merely undefined.

. . . .

Sound philosophy will not authorise me to alter this opinion of the mortality of man on earth till it can be clearly proved that the human race has made, and is making, a decided progress towards an illimitable extent of life. And the chief reason why I adduce the two particular instances from animals and plants was to expose and illustrate, if I could, the fallacy of that argument which infers an unlimited progress merely because some partial improvement has taken place, and that the limit of this improvement cannot be precisely ascertained.

The capacity of improvement in plants and animals, to a certain degree, no person can possibly doubt. A clear and decided progress has already been made; and yet I think it appears that it would be highly absurd to say that this progress has no limits. In human life, though there are great variations from different causes, it may be doubted whether since the world began any organic improvement whatever of the human frame can be clearly ascertained. The foundations, therefore, on which the arguments for the organic perfectibility of man rest are unusually weak, and can only be considered as mere conjectures. It does not however by any means seem impossible that by an attention to breed a certain degree of improvement similar to that among animals might take place among men. Whether intellect could be communicated may be a matter of doubt; but size, strength, beauty, complexion, and perhaps even longevity, are in a degree transmissible. The error does not lie in supposing a small degree of improvement impossible, but in not discriminating between a small improvement, the limit of which is undefined and an improvement really unlimited. As the human race however could not be improved in this way without condemning all the bad specimens to celibacy, it is not probable that an attention to breed should ever become general; indeed I know of no well-directed attempts of this kind, except in the ancient family of the Bickerstaffs, who are said to have been very successful in whitening the skins and increasing the height of their race by prudent marriages, particularly by that very judicious cross with Maud the milkmaid, by which some capital defects in the constitutions of the family were corrected.

Just how much Charles Darwin owed to his grandfather, Erasmus, is a matter of dispute. He himself thought it was very little. Several of his biographers, however, are inclined to believe that he owed more to his grandfather than he realized. At any rate the historian of biology will find that the writings of Erasmus Darwin will repay an intimate study. Here we will quote only a single passage from his *Temple of Nature*, London, 1803 wherein he describes the struggle for existence in no uncertain language:

Fell Aestrus buries in her rapid course
Her countless brood in stag, or bull, or horse;
Whose hungry larva eats its living way,
Hatch'd by the warmth, and issues into day.
The wing'd Ichneumon for her embryon young
Gores with sharp horn the caterpillar throng.
The cruel larva mines its silky course,
And tears the vitals of its fostering nurse.
While fierce Libellula with jaws of steel
Ingulfs an insect-province at a meal;
Contending bee-swarms rise on rustling wings,
And slay their thousands with envenom'd stings.

Yes! smiling Flora drives her armed car
Through the thick ranks of vegetable war;
Herb, shrub, and tree, with strong emotions rise
For light and air, and battle in the skies;
While roots diverging with opposing toil

Contend below for moisture and for soil;
Round the tall Elm the flattering Ivies bend,
And strangle, as they clasp their struggling friend;
Envenom'd dews from Mancinella flow
And scald with caustic touch the tribes below;
Dense shadowy leaves on stems aspiring borne
With blight and mildew thin the realms of corn;
And insect hordes with restless tooth devour
The unfolded bud, and pierce the ravell'd flower

In ocean's pearly haunts, the waves beneath
Sits the grim monarch of insatiate Death;
The shark rapacious with descending blow
Darts on the scaly brood, that swims below;
The crawling crocodiles, beneath that move,
Arrest with rising jaws the tribes above;
With monstrous gape sepulchral whales devour
Shoals at a gulp, a million in an hour.
—Air, earth, and ocean, to astonish'd day
One scent of blood, one mighty tomb display!
From Hunger's arms the shafts of Death are hurl'd,
And one great Slaughter-house the warring world!

The brow of Man erect, with thought elate,
Ducts to the mandate of resistless fate;
Nor Love retains him, nor can Virtue save
Her sages, saints, or heroes from the grave.
While cold and hunger by defect oppress,
Repletion, heat, and labour by excess,
The whip, the sting, the spur, the fiery brand,
And, cursed Slavery! thy iron hand;
And led by Luxery Disease's trains,
Load human life with unextinguish'd pains.

Malthus' *Essay* was written within two years of the end of the eighteenth century. At the beginning of the nineteenth century evolution was very definitely "in the air". Perhaps it would be more accurate to describe the belief in evolution at this time as underground, as held by many naturalists but submerged by the dominant belief in the fixity of species. Kohlbrugge [8] has shown how many biologists were ready to welcome the doctrine of evolution when it emerged as a respectable scientific hypothesis. Of all these evolutionists Prichard came closest to explaining the origin of new forms through the operation of natural selection although he never actually stated the proposition in so many words. His great anthropological work started out as a dissertation, *De humani generis varietate* (1808). This was translated into English and published as *Researches into the Physical History of Man*, London, 1813. In the second edition, issued in 1826, Prichard's views on evolution reached their maximum expression. In subsequent editions he showed more and more interest in the minutiae of his subject and he seemed to forget or become indifferent to its broader evolutionary aspects.

---

[8] Kohlbrugge, J. H. F. "War Darwin ein originelles Genie?" *Biol. Centralblatt.* **35**: 93–11 (1915).

Prichard's conception of evolution seems very modern. He distinguished very clearly between congenital and acquired characteristics and denied the inheritance of the latter. He recognized the sudden appearance of new forms in domestic stock which were established by artificial selection when they proved to be valuable. He considered that new races of man were formed in a manner analogous to the formation of new varieties of animals and plants and he described how a race, well fitted to one climate could not compete with an indigenous, well adapted race when it moved to a different climate. He even noted the evidence for evolution which existed in animal distribution, and in the fitness of each variety for its particular habitat. Indeed we are justified in marveling how Prichard avoided discovering that natural selection was a major factor in organic evolution. The following passages are from the second (1826) edition. From p. 573:

> But there are more important differences in the constitution of the European and the Negro, by which these varieties of our species seem to be adapted respectively to the countries which they inhabit.
>
> It seems that people descended from the European nations bear with the difficulty an abode in Africa between the tropics. The insalubrity of the intertropical African climate to the constitution of Europeans is extreme, and has hitherto been sufficient, notwithstanding innumerable attempts at colonization, to prevent the growth of any white population of the African coasts.
>
> . . . .
>
> On the other hand, Negroes seem under a physical disability to establish themselves in Europe and other northern countries, otherwise we should find them here in numbers. They are here, more than any other class of people, subject to phthisical and scrofulous complaints, and are seldom, under the most favourable circumstances, healthy. The diseases to which both kinds of people are subject in the climate appropriated to the other, is the impediment which has prevented large colonies of whites from forming themselves and multiplying in tropical

Africa, and of Negroes in the North. We are told indeed by Herodotus, that there was once founded at Colchis a colony of African blacks, but they have long since dwindled away and disappeared.

It appears to result from the foregoing facts, to which a great many others might probably be added, that in mankind, as in some other races, particular varieties are adapted by constitution and physical peculiarities to particular local situations.

These remarks, if they are well founded, serve to illustrate the doctrine of variation, or deviation, in the races of animals in general, and they seem to lead us to the conclusion, that this is not merely an accidental phænomenon, but a part of the provision of nature for furnishing to each region an appropriate stock of inhabitants, or for modifying the structure and constitution of species, in such a way as to produce races fitted for each mode and condition of existence. A great part of this plan of local adaptation appears to have been accomplished by the original modification of a genus into a variety of species. It has been further continued, and the same end promoted, by the ramification of a species into several varieties.

From p. 581:

A question now presents itself; how these varieties are developed and preserved in connexion with particular climates and differences of local situation?

One cause which tends to maintain this relation is obvious. Individuals and families, and even whole colonies, perish and disappear in climates for which they are, by peculiarity of constitution, not adapted. Of this fact proofs have been already mentioned.

Besides, it appears probable that those local circumstances, which are most congenial to particular races, do in fact promote the appearance of those varieties which are best suited to them, or tend to give rise to their production in the breed.

A contemporary of Prichard's, William Charles Wells, was the first who explained organic evolution by natural selection. Wells, a doctor of medicine, was born in Charleston, South Carolina, but removed to England at the outbreak of the Revolutionary War. Darwin did not know of Wells' work and when he learned of it, he minimized it somewhat, stating that Wells applied natural selection "only to the races of man, and to certain characters alone". Wells did illustrate the principle by applying it to human skin color but he obviously understood its wider application and stated, "Similar facts occur in respect to other species of animals." Wells' theory, however, received little attention, although he certainly did not hide it under a bushel. In 1813, he presented it before the leading scientific society of his time, the Royal Society of London, and it was included in his famous post-humous, *Two Essays upon Double and Single Vision, etc.* London, 1818. The relevant part of Wells' description of natural selection is here quoted in full:

> On considering the difference of colour between Europeans and Africans a view has occurred to me of this subject, which has not been given by any author, whose works have fallen into my hands. I shall, therefore, venture to mention it here, though at the hazard of its being thought rather fanciful than just.
>
> There is no circumstance, perhaps, in which these two races differ so much, as in their capacity to bear, with impunity, the action of the causes of many diseases. The fatality to Europeans of the climate of the middle parts of Africa, which are, however, inhabited by negroes without injury to their health, is well known. Let it then be supposed, that any number of Europeans were to be sent to that country, and that they were to subsist themselves, by their bodily labour; it seems certain, that the whole colony would soon become extinct. On the other hand, the greater liability of negroes in Europe to be attacked with fatal diseases is equally well established. If, therefore, a colony of the latter race were brought to Europe, and forced to labour in the open air for their subsistence, many of them would quickly die, and the remainder, from their inability to make great bodily exertions in cold weather, and their being frequently diseased, would be prevented from working an equal number of days in the year with the whites. The consequence would be, that without taking farther into account the unfriendliness of the climate to them, their gains would be inadequate to the maintenance of themselves and their families. They would thence become feeble, and be rendered still more incapable of supporting life by their labour. In the meantime, their children would die from want, or disease induced by deficient or improper nourishment, and in this way, a colony of the negro race in a cold country would quickly cease to exist.
>
> This difference in the capacity of the two races to resist the operation of the causes of many diseases, I assume as a fact, though I am utterly unable to explain it. I do not, however, suppose that their different susceptibility of diseases depends, properly, on their difference of colour. On the contrary, I think it probable that this is only a sign of some difference in them, which, though strongly manifested by its effects in life, is yet too subtle to be discovered by an anatomist after death; in like manner as a human body, which is incapable of receiving the small-pox, differs in no observable thing from another, which is still liable to be affected with that disease.
>
> Regarding then as certain, that the negro race are better fitted to resist the attacks of the diseases of hot climates than the white, it is reasonable to infer, that those, who only approach the black race, will be likewise better fitted to do so, than others who are entirely white. This is, in fact, found to be true, with regard to the mixture of the two races; since mulattoes are much more healthy in hot climates than whites.

But amongst men, as well as among other animals, varieties of a greater or less magnitude are constantly occurring. In a civilized country, which has been long peopled, those varieties, for the most part, quickly disappear from the inter-marriages of different families. Thus, if a very tall man be produced, he very commonly marries a woman much less than himself, and their progeny scarcely differs in size from their countrymen. In districts, however, of very small extent, and having little intercourse with other countries, an accidental difference in the appearance of the inhabitants will often descend to their late posterity. The clan of the Macras, for instance, possess both sides of Loch-Duich in Scotland; but those who inhabit one side of the loch are called the black Macras, and the others the white, from a difference which has always been observed in their complexions. Again, those who attend to the improvement of domestic animals, when they find individuals possessing, in a greater degree than common, the qualities they desire, couple a male and female of these together, then take the best of their offspring as a new stock, and in this way proceed, till they approach as near the point in view, as the nature of things will permit. But what is here done by art, seems to be done, with equal efficacy, though more slowly, by nature, in the formation of varieties of man, which would occur among the first few and scattered inhabitants of the middle regions of Africa some one would be better fitted than the others to bear the diseases of the country. This race would consequently multiply, while the others would decrease, not only from their inability to sustain the attacks of disease, but from their incapacity of contending with their more vigorous neighbors. The colour of this vigorous race I take for granted, from what has been already said, would be dark. But the same disposition to form varieties still existing, a darker and a darker race would in the course of time occur, and as the darkest would be best fitted for the climate, this would at length become the most prevalent, if not the only race, in the particular country in which it had originated.

In like manner, that part of the original stock of the human race, which proceeded to the colder regions of the earth, would in process of time become white, if they were not originally so, from persons of this colour being better fitted to resist the diseases of such climates, than others of a dark skin.

The cause which I have stated, as likely to have influence on the colour of the human race, would necessarily operate chiefly during its infancy, when a few wandering savages, from ignorance and improvidence, must have found it difficult to subsist throughout the various seasons of the year, even in countries the most favourable to their health. But, when men have acquired the knowledge of agriculture, and other arts, and in consequence adopt a more refined mode of life, it has been found, that an adherence to their ancient customs and practices will preserve them long as a distinct race from the original inhabitants of the country to which they had emigrated. Examples of this kind are frequent in the islands in the eastern seas in the torrid zone, where the inhabitants of the sea-coasts, evidently strangers, are in some degree

> polished, and of a brown colour, while the ancient natives, who live in the interior parts, are savage and black. Similar facts occur in respect to other species of animals. It seems certain, for instance, that fine woolled sheep, like the Spanish, never both arose and sustained their breed in the northern parts of Europe; yet, by care, this feeble race, after being formed in Spain, has been propagated and preserved in very cold countries. Thus the late Mr. Drylander, the learned librarian of the Royal Society, informed me, that the breed of fine woolled Spanish sheep had been kept perfect in Sweden during a very long term of years, I think he said a century. If, then, my memory be accurate upon this point, we have here an example of a variety of animals, much more liable to be affected by external circumstances than the human race, being preserved without change, in a country very different from their own, by assimilating their new state as much as possible to their old, during at least fifty generations, that is, during a period equivalent to 1500 years in the history of man.

William Lawrence, like Prichard, almost arrived at an explanation of evolution through the action of natural selection and, like Prichard, just failed to make the ultimate logical inference. In his *Lectures on the Natural history of Man*, London, 1819, he discussed the inheritance of acquired characters but discarded the idea. On the other hand he saw how geographical isolation could establish new varieties. He recognized the role of sexual selection, and the function of artificial selection in general. He even hoped that a logical program of selective breeding would produce an eugenic improvement in rundown human stock. How close he came to the concept of natural selection is shown by the juxtaposition of the following citations. After rejecting the inheritance of acquired characters, he stated (p. 348):

> It is obvious that the external influences just considered, even though we should not allow them to a much greater influence on individuals than experience warrants us in admitting, would be still entirely inadequate to account for those signal diversities which constitute differences of race in animals. These can be explained only by two principles already mentioned; namely, the occasional production of an offspring with different characters from those of the parents, as a native or congenital variety; and the propagation of such varieties by generation.

Artificial selection is described (p. 179):

> The greatest differences are produced when man regulates the sexual intercourse of animals; by selecting individuals to breed from, he can effect the most surprising changes in form and qualities, as the examples of the pig, sheep, horse, cow, and dog, will abundantly evince. The deviation has become at last so great, that the original stock from which the animals descended is doubtful.

Selection is applied to man (p. 313):

> The hereditary transmission of physical and moral qualities, so well understood and familiarly acted on in the domestic animals, is equally true of man. A superior breed of human beings could only be produced by selections and exclusions similar to those so successfully employed in rearing our more valuable animals.

Finally he noted that the differences which exist between human races in a state of nature were produced by a process analogous to artificial selection (p. 375):

> The facts and observations adduced in this section lead us manifestly to the following conclusions; 1st. that the differences of physical organization and of moral and intellectual qualities, which characterize the several races of our species, are analogous in kind and degree to those which distinguish the breeds of the domestic animals, and must, therefore, be accounted for on the same principles. 2dly, that they are first produced, in both instances, as native or congenital varieties, and then transmitted to the offspring in hereditary succession.

The results of sexual selection are described (p. 309):

> The great and noble have generally had it more in their power than others to select the beauty of nations in marriage; and thus, while, without system or design, they gratified merely their own taste, they have distinguished their order, as much by elegant proportions of person, and beautiful features, as by its prerogatives in society. "The same superiority," says Cook, "which is observable in the dress of nobles in all the other islands, is found here (Sandwich Islands). Those whom we saw, were, without exception, perfectly well formed; whereas, the lower sort, beside their general inferiority, are subject to all the variety of make and figure that is seen in the populace of other countries."
>
> In no instance, perhaps, has the personal beauty of a people been more improved, by introducing handsome individuals to breed from, than in the Persians, of whom the nobility have, by this means, completely succeeded in washing out the stain of their Mongolian origin.

Patrick Matthews' claim to priority over Darwin and Wallace in originating the idea of natural selection was acknowledged by Darwin in the preface to the second edition of the *Origin of Species.* Matthews' claim was based on his book, *Naval timber and aboriculture*, London, 1831, for he had expressed his belief in an appendix that the present species had not been created as they were but that they had evolved from other forms. No one seems to have noticed the idea until after Darwin's work was published, when Matthew called attention to his earlier statement of natural selection in an article printed in the *Gardner's Chronicle,* on April 7, 1860. Darwin's reply to Matthews' claim was to acknowledge Matthews' priority completely and to offer in explanation the fact that neither he nor apparently any other naturalist had ever heard of Matthews' views.

The following quotations are from the passages republished by Matthews in the *Gardner's Chronicle.*

> There is a law universal in Nature, tending to render every reproductive being the best possibly suited to its condition that its kind, or that organized matter is susceptible of, which appears intended to model the physical and mental or instinctive powers, to their highest perfection, and to continue them so. This law sustains the lion in his strength, the hare in her swiftness, and the fox in his wiles. As Nature, in all her modifications of life, has a power to increase far beyond what is needed to supply the place of what falls by Time's decay, those individuals who possess not the requisite strength, swiftness, hardiness, or cunning, fall prematurely without reproducing—either a prey to their natural devourers, or sinking under disease, generally induced by want of nourishment, their place being occupied by the more perfect of their own kind, who are pressing on the means of subsistence.

. . . .

> The self-regulating adaptive disposition of organised life may, in part, be traced to the extreme fecundity of nature, who, as before stated, has, in all the varieties of her offspring, a prolific power much beyond (in many cases a thousandfold) what is necessary to fill up the vacancies caused by senile decay. As the field of existence is limited and pre-occupied, it is only the hardier, more robust, better suited to circumstance individuals who are able to struggle forward to maturity, these inhabiting only the situations to which they have superior adaptation and greater power of occupancy than any other kind; the weaker, less circumstance-suited being prematurely destroyed. This principle is in constant action, it regulates the colour, the figure, the capacities, and instincts; those individuals of each species, whose colour and covering are best suited to concealment or protection from enemies, in defense from vicissitude and inclemencies of climate, whose figure is best accommodated to health, strength, defense, and support; whose capacities and instincts can best regulate the physical energies to self-advantage according to circumstances in such immense waste of primary and youthful life, those only come forward to maturity from the strict ordeal by which Nature tests their adaptation to her standard of perfection and fitness to continue their kind by reproduction.
>
> From the unremitting operation of this law acting in concert with the tendency which the progeny have to take the more particular qualities of the parents, together with the connected sexual system of vegetable, and instinctive limitation to its own kind in animals, a considerable

> uniformity of figure, colour, and character, is induced, constituting species; the breed gradually acquiring the very best possible adaptation of these to its condition which it is susceptible of, and when alteration of circumstances occurs, thus changing in character to suit these as far as its nature is susceptible of change.

After the quotations from the writings of Prichard, Wells, Lawrence and Matthews, the short but fluent description of the struggle for existence by August Pyrimus De Candolle is somewhat of an anti-climax. The passage, however, was quoted by Sir Charles Lyell and it obviously influenced Charles Darwin himself. It occurs in De Candolle's *Physiologie Végétal*, Paris, 1832, in a chapter devoted to parasitism. From Vol. III, p. 1401:

> All the plants of a given country are at war one with another. The first which establish themselves by chance in a particular spot, tend, by the mere occupancy of space, to exclude other species—the greater choke the smaller, the longest livers replace those which last for a shorter period, the more prolific gradually make themselves masters of the ground, which species multiplying more slowly would otherwise fill.

Lyell used this passage to introduce a chapter on the balance of nature in his *Principles of Geology*, London, 1833. This chapter deals in some detail with the intense competition between species and with the complex interrelationship between various predators and the organisms on which they prey. It describes the reproductive potentialities of plants, insects, fish and mammals and shows how population pressure operates in nature. Lyell also described the struggle for existence in its inter-species phase but he did not mention variation and the intense intra-species competition. Darwin refers to the description of the struggle by De Candolle in his MSS of 1842 and 1844 and couples the names of De Candolle and Lyell when he mentions it in the *Origin*.

Isidore Geoffrey St. Hilaire, well known for his advocacy of evolution, noted briefly that nature would preserve certain variations and eliminate others in adapting an organism to a changed environment. Such an evolutionary modification would occur, of course, in the absence of population pressure and without any struggle for existence. The statement in question is in an article, "Influence du monde ambient pour modifier les formes animals," *Mem. de l'Acad. de Sci.*, **12**: 79, 1833:

> I believe breathing constitutes an arrangement so powerful for the disposition of animal forms, that it is not in the least necessary for the

medium of the respiratory fluids to be modified suddenly and strongly, to cause forms to be altered discriminately. The slow action of time ordinarily provides this, particularly if the organism survives an immediate disaster. The indiscernible modifications from one century to another become cumulative and add up to such a result, that breathing may become a difficult task and finally impossible for certain systems of organs: it then needs and creates for itself another arrangement, perfecting or altering the pulmonary cells, in which it operates; happy or lethal modifications occur, which propagate themselves and which influence the rest of the animal economy in each detail. For if these modifications lead to dangerous effects, the animal which experiences them ceases to exist, to be replaced by others, with forms a bit changed, and changed to suit the new circumstances.

If we consider only those naturalists who almost arrived at a conception of natural selection, together with Wells and Matthews who did reach it, we would have a very false picture of the scientific outlook of the early nineteenth century. The works cited in this paper must have been read by at least a few intelligent people. There was also a widespread, if underground, belief in the instability of species, but progress toward a comprehensive theory of organic evolution just did not occur. The dominant thought of the period indorsed a planned universe with a special creation for each species. The widespread acceptance of teleology made natural selection an unneeded hypothesis and probably served as a protective inoculation against the spread of the doctrine. In the leading history of science written at this time, William Whewell's *History of the Inductive Sciences*, no mention whatever is made of natural selection. Thomas Huxley called attention to this fact twice. "It is interesting to observe that the possibility of a fifth alternative [natural selection] in addition to the four he has stated, has not dawned upon Dr. Whewell's mind." (*Life and Letters of Chas. Darwin*, II: 195.) "Whewell had not the slightest suspicion of Darwin's main theorem, even as a logical possibility." (*Science* supplement to *The Reign of Queen Victoria.*) Romanes[9] has shown, however, that Whewell did contemplate natural selection in his *Astronomy and General Physics*, London, 1833, but rejected it violently in favor of teleology. From pp. 24–26, Ed. of 1852, the italics are the writer's:

But in the existing state of things, the duration of the earth's revolution round the sun, and the duration of the revolution of the vegetable functions of most plants are equal. These two periods are *adjusted* to each other. The stimulants which the elements apply come at such intervals, and continue for such time that the plant is supported in health and vigour, and enabled to reproduce its kind. Just such a portion of time is measured out for the vegetable powers to execute their task, as enables them to do so in the best manner.

---

[9] Romanes, George John, *Darwin and after Darwin*, Chicago, 1892.

Now such an adjustment must surely be accepted as a proof of design exercised in the formation of the world. Why should the solar year be so long and no longer? or, this being of such a length, why would the vegetable cycle be exactly of the same length? Can this be chance? And this occurs, it is to be observed, not in one, or in a few species of plants, but in thousands. Take a small portion only of known species as the most obviously endowed with this adjustment, and say ten thousand. How should all these organised bodies be constructed for the same period of a year? How should all these machines be wound up so as to go for the same time? Even allowing that they could bear a year of a month longer or shorter, how do they all come within such limits? No chance could produce such a result. And if not by chance, how otherwise could such a coincidence occur, than by an intentional adjustment of these two things to one another? by a selection of such an organisation in plants, as would fit them to the earth on which they were to grow; by an adaptation of construction to conditions; of the scale of the conditions.

It cannot be accepted as an explanation of this fact in the economy of plants, that it is necessary to their existence; that no plants could possibly have subsisted, and come down to us, except those which were thus suited to their place on the earth. This is true; but this does not at all remove the necessity of recurring to design as the origin of the construction by which the existence and continuance of plants is made possible. A watch could not go, except there were the most exact adjustment in the forms and positions of its wheels; yet no one would accept it as an explanation of the origin of such forms and positions, that the watch would not go if these were other than they are. *If the objector were to suppose that plants were originally fitted to years of various lengths, and that such only have survived to the present time, as had a cycle of a length equal to our present year, or one which could be accommodated to it; we should reply, that the assumption is too gratuitous and extravagant to require much consideration; but that, moreover, it does not remove the difficulty.* How came the functions of plants to be *periodical* at all? Here is, in the first instance, an agreement in the form of the laws that prevail in the organic and in the inorganic world, which appears to us a clear evidence of design in their Author. And the same kind of reply might be made to any similar objection to our argument. *Any supposition that the universe has gradually approximated to that state of harmony among the operations of its different parts, of which we have one instance in the coincidence now under consideration, would make it necessary for the objector to assume a previous state of things preparatory to this perfect correspondence.* And in this preparatory condition we should still be able to trace the rudiments of that harmony, for which it was proposed to account: so that even the most unbounded license of hypothesis would not enable the opponent to obliterate the traces of an intentional adaptation of one part of nature to another.

It is an interesting commentary on Whewell's character that, although he had known Darwin personally for many years, he would not allow a copy of the *Origin of Species* in the Library of Trinity College, Cambridge.[10]

Dean William Herbert, known for his work on plant hybridization, is quoted by Darwin, because of his recognition of the struggle for existence. In an article on bulbous plants he approached very closely to the natural selection hypothesis when he suggested that winter hardiness might become established in a hybrid stock through the survival of chance variations. From: *Amaryllidaceae; etc.* London, 1837, p. 347.

> . . . for it does not appear that in reality any plant becomes acclimated under our observation, except by crossing with a hardier variety, or by the accidental alteration of constitution in some particular seedling; not that any period of time does in fact work an alteration in the constitution of an individual plant, so as to make it endure a climate which it was originally unable to bear. . . .

In the same year that Herbert's *Amaryllidaceae* was published Darwin started to incubate his great idea. He had not as yet read Malthus but his *Note Book* of 1837 contains a passage which indicates that he saw then how nature could select adaptation or fitness. "With respect to extinction, we can easily see that a variety of the ostrich (Petise), may not be well adapted, and thus perish out; or on the other hand, like Orpheus, being favourable, many might be produced. This requires the principle that the permanent variations produced by confined breeding and changing circumstances are continued and produce(d) according to the adaptation of such circumstances, and therefore that death of species is a consequence (contrary to what would appear in America) of non-adaptation of circumstances." In October of the next year he read the famous *Essay of the Principles of Population* and, as he recorded in the *Autobiography*, found the clue which enabled him to identify the agency in nature which could select new variations and form new species, "being well prepared to appreciate the struggle for existence . . ., it at once struck me that under these circumstances favourable variations would tend to be preserved, and unfavourable ones to be destroyed. The result of this would be the formation of new species. Here then I had at last got a theory by which to work."

---

[10] West, Geoffrey, *Charles Darwin, the Fragmentary Man*, London, 1937.

Darwin now had a clear conception of natural selection as the essential cause of the origin of species and of organic evolution in general. In 1842, he made a crude outline of his hypothesis which was developed into a more finished essay in 1844. The contents of these two MSS are now readily available.[11] It was an abstract of this essay of 1844, together with a letter on natural selection written by Darwin in 1857 to Asa Gray, which were read before the Linnean Society on July 1, 1858 as Darwin's contribution to the famous Darwin-Wallace session. While Darwin discussed his idea of natural selection with his friends, some of whom were the leading scientists of the time, he did not publish it until twenty years after its inception. Meanwhile, Sir Richard Owen, Herbert Spencer and Charles Naudin came very close to the realization of the existence of natural selection and its logical implications and Alfred Russel Wallace actually reached Darwin's conclusions independently.

Perhaps nothing in the history of natural selection is as obscure as the role played by Sir Richard Owen. Owen was a patronizing friend of Darwin's when the latter was a young unknown naturalist, a bitter and vindictive enemy when Darwin became a famous scientist. The concensus of opinion seems to be that Owen was a brilliant but unscrupulous man, jealous and unforgiving. Many of the anonymous attacks on Darwin which followed the publication of the *Origin of Species* were composed by Owen and it is not to be wondered at that Darwin could not look upon Owen with his usual objectivity, particularly when Owen laid claim to the theory of natural selection. Darwin confessed to being completely puzzled by Owen's statement that he (Owen) had anticipated the Darwin-Wallace hypothesis. Actually the records of the case are clear and the facts can be ascertained easily.

Before we can evaluate the various charges and counter-charges in the Darwin-Owen dispute we must examine it from the viewpoint of the mid-nineteenth century. Today we accept the fact of evolution as a matter of course and are interested in such hypotheses as the inheritance of acquired character or natural selection only as they explain how evolution occurred. A hundred years ago the case was very different. Then the belief in the inheritance of acquired characters, which had been accepted almost universally for over two thousand years, was held by a number of scientists (Lyell for example) who did not believe in evolution. Likewise in 1850, the date of Owen's contribution, scientists could believe that natural selection was a potent force yet refuse very logically to credit it with the power of producing new species. If we keep these facts in mind the case becomes simple.

---

[11] Darwin, Chas., *The Foundation of the Origin of Species* (Edited by Francis Darwin), Cambridge, 1909.

Darwin's controversial writings shed little light on the matter. The following quotation is from the historical sketch which appeared in the later editions of the *Origin:*

> When the first edition of this work was published, I was so completely deceived, as were many others, by such expressions as "the continuous operation of creative power", that I included Prof. Owen with other paleontologists as being firmly convinced of the immutability of species; but it appears (*Anat. of Vertebrates,* III: 796) that this was on my part a preposterous error. In the last edition of this work I inferred, and the inference still seems to me perfectly just, from a passage beginning with the words "no doubt the type form": &c. (*Ibid.,* I: xxxv) that Professor Owen admitted that natural selection may have done something in the formation of a new species; but this it appears (*Ibid.,* III: 798) is inaccurate and without evidence. I also gave some extracts from a correspondence between Professor Owen and the Editor of the *London Review* from which it appeared manifest to the Editor as well as to myself, that Professor Owen claimed to have promulgated the theory of natural selection before I had done so; and I expressed my surprise and satisfaction at this announcement; but as far as it is possible to understand certain recently published passages (*Ibid.,* III: 798) I have either partially or wholly again fallen into error. It is consolatory to me that others find Professor Owen's controversial writings as difficult to understand and to reconcile with each other, as I do.

The controversy started in 1866 when Owen published the first two volumes of the *Comparative Anatomy and Physiology of Vertebrates.* In the preface he quoted a paragraph from an article "On the Genus Dinornis", which he had published in *Zool. Trans.* (**4**: 15) in 1850. The paragraph follows:

> In proportion to its bulk is the difficulty of the contest which, as a living organised whole, the individual of such species has to maintain against the surrounding agencies that are ever tending to dissolve the vital bond and subjugate the living matter to the ordinary chemical and physical forces. Any changes, therefore, in such external agencies as a species may have been originally adapted to exist in, will militate

> against that existence in a degree proportionate, perhaps in a geometrical ratio, to the bulk of the species. If a dry season be gradually prolonged, the large mammal will suffer from the drought sooner than the small one; if such alteration of climate affect the quantity of vegetable food, the bulky Herbivore will first feel the effects of stinted nourishment; if new enemies are introduced, the large and conspicuous quadruped or bird will fall a prey, whilst the smaller species conceal themselves and escape. Smaller animals are usually, also, more prolific than larger ones.

Here Owen visualized natural selection as eliminating species but not as creating new ones. Owen continued:

> The actual presence, therefore, of small species of animals in countries where larger species of the same natural families formerly existed, is not the consequence of any gradual diminution of the size of such species, but is the result of circumstances which may be illustrated by the fable of the "Oak and the Reed"; the smaller and feebler animals have bent and accommodated themselves to changes which have destroyed the larger species. They have fared better in the "battle of life".
>
> Accepting this explanation of the extirpation of species as true, Mr. Wallace (*Proc. Linnean Soc.*, Aug. 1858, p. 57) has applied it to the extirpation of varieties; and as these do arise in a wild species, he shows how such deviations from type may either tend to the destruction of a variety, or to adapt a variety to some changes in surrounding conditions, under which it is better calculated to exist, than the type-form from which it deviated.
>
> . . . .
>
> In his work *"On the Origin of Species by Natural Selection"*, Mr. Darwin more fully exemplifies, conjecturally, the reciprocal influence of external conditions and inherent tendencies to variety, in carrying on, as he believes, the deviations from type to specific and higher degrees of difference.

Owen next explained why he was unable to agree with Wallace and Darwin and stated that their theories were powerless to explain the great differences between species. In the third volume, which appeared in 1868 (Footnote, p. 798), Owen describes the origin of the controversy, states his claim of priority to the concept of natural selection and sticks his famous stiletto into Darwin. Incidentally he demonstrates his own peculiar intellectual limitations:

> A critic of the first volume of the present work, switching over the pages of the "Preface" with the speed they merited at his hands, caught sight of the words, "contest of existence", "battle of life"; and thereupon dashed off with—"We would call attention to the following passage, and ask whether it is not actually an admission of the Darwinian Theory!"

("*London Review*", April 28, 1866, p. 483); then pastes in the slip beginning with "the actual presence" to "fared better in the battle of life". With the bulk of two volumes before him, an able reviewer could hardly be expected to waste valuable time upon "notes", and so the fact escaped him that the "admission" or "adoption" was, in whatever degree it might relate to the D. T., an anticipation.

Oddly enough, another reviewer (if haply the same meritorious labourer may not have been doing this sort of work for both periodicals) makes the same transposition of dates, mistaking a quotation for text; *e.g.* "not the least important feature in the work before us is, that it contains a partial concurrence, on the part of the author, in the theory of *Natural Selection*". And the same cutting does duty as the "piece justificative", viz. "The actual presence", &c. to "battle of life".—("*Popular Science Review*", April, 1866, p. 212).

Having regard to intelligent countrymen and countrywomen taking scientific sustenance from these weekly and monthly sources, and who might never see the pages of the work reviewed, I ventured to call attention to the omitted reference in the footnote of my "Preface" viz. to the volume of "*Transactions of the Zoological Society*", 1850, in which my theory of the extinction and conservation of species appeared, including the passage quoted, with the obvious remark, that "if the difference between 1858 (date of the D. T. or 'Natural Selection') and 1868 (date of Vol. 1 of *Anat. of Vertebrates*) puts the writer of the latter date in the subordinate relation of 'admitter' or 'adopter'—tacit or otherwise—to the author of the same theory at the earlier date, the writer of 1858 must stand in the same relation to the author of the same theory of 1850." (Letter to Ed. of "*London Review*", May 1st, 1866.)

Of course, to every competent judge, the difference between a theory founded on the application of the principle of the contest for existence to the preservation or extinction of certain species, and that of the theory of the origin of all species partially based upon the same principle, must have been obvious; nor was any pretention advanced, in the letter rectifying the date of the "idea", to the ample and instructive degree in which it had been worked out, and doubtless as an original thought, by the accomplished author of the *Origin of Species.*

Darwin's relations were much more pleasant with his next partial anticipator, Herbert Spencer, who soon became an enthusiastic Darwinian. Spencer called attention in *The Principles of Biology,* London, 1874, to a paragraph he had published in the *Westminister Reviews* in April 1852, only to show how near he could come to a great generalization and yet not see it. The following quotations are taken from the reprint, *A New Theory of Population* by R. T. Trall, New York, 1853. Its debt to Malthus is obvious. From pp. 40–42:

For, as we all know, this excess of fertility entails a constant pressure of population upon the means of subsistence; and, as long as it exists, must continue to do this. Looking only at the present and the immediate future, it is unquestionably true that if unchecked, the rate of increase of people would exceed the rate of increase of food. It is clear that the wants of their redundant numbers constitute the only stimulus mankind have to a greater production of the necessaries of life; for, were not the demand beyond the supply, there would be no motive to increase the supply. Moreover, this excess of demand over supply, and this pressure of population, of which it is the index, cannot be eluded. Though by the emigration that takes place when the pressure arrives at a certain intensity, a partial and temporary relief may be obtained, yet, as by this process all habitable countries must gradually become peopled, it follows, that in the end the pressure, whatever it may be, must be borne in full.

But this inevitable redundancy of numbers—this constant increase of people beyond the means of subsistence—involving as it does an increasing stimulus to better the modes of producing food and other necessaries—involves also an increasing demand for skill, intelligence, and self-control—involves, therefore, a constant exercise of these, that is —involves a gradual growth of them. Every improvement is at once the product of a higher form of humanity, and demands that higher form of humanity to carry it into practice.

. . . .

In all cases, increase of numbers is the efficient cause. Were it not for the competition this entails, more thought would not daily be brought to bear upon the business of life; greater activity of mind would not be called for; and development of mental power would not take place.

. . . .

And here it must be remarked, that the effect of pressure of population, in increasing the ability to maintain life, and decreasing the ability to multiply, is not a uniform effect, but an average one. In this case, as in many others, Nature secures each step in advance by a succession of trials, which are perpetually repeated, and cannot fail to be repeated, until success is achieved. All mankind in turn subject themselves more or less to the discipline described; they either may or may not advance under it; but, in the nature of things only those who do advance under it eventually survive. For necessarily, families and races whom this increasing difficulty of getting a living which excess of fertility entails, does not stimulate to improvements in production—that is, to greater mental activity—are on the high road to extinction; and must ultimately be supplanted by those whom the pressure does so stimulate. This truth we have recently seen exemplified in Ireland. And here, indeed, without further illustration, it will be seen that premature death, under all its forms, and from all its causes, cannot fail to work in the same direction. For as those prematurely carried off must, in the average of cases, be those in whom the power of self-preservation is the least, it unavoidable follows that those left behind to continue the race are those

> in whom the power of self preservation is the greatest—are the select of their generation. So that, whether the dangers to existence be of the kind produced by excess of fertility, or of any other kind, it is clear, that by the ceaseless exercise of the faculties needed to contend with them, and by the death of all men who fail to contend with them successfully, there is insured a constant progress toward a higher degree of skill, intelligence, and self-regulation—a better coordination of actions—a more complete life.

The same year (1852) that Spencer approximated a description of natural selection Charles Victor Naudin in the *Revue Horticole* (pp. 103–105) came even closer to forestalling Darwin. Obviously so many biologists were approaching the conception of natural selection that it was only a question of time until it would break out into the open. After describing artificial selection, Naudin continues:

> We don't believe that nature has proceeded to form its species, in another manner than we ourselves proceed in creating our varieties; let us say even better: it is her very procedure that we have transported to our practice. We draw forth a variety which will correspond to such of our needs, and we choose among the great number of individuals of this species, to make from it the point of departure for a new line, those which appear to us to separate themselves already from the specific type in the sense which suits us, and by a rational sorting and following the products obtained, we arrive, at the end of an indetermined number of generations, at the creation of varieties or artificial species which respond more or less well to the ideal type which we have formed, and which transmits so much better to their descendents the acquired characteristics which our efforts have borne on a greater number of generations. Such is, in our ideas, the means followed by nature; like us, she wanted to form races to appropriate them to her needs; and, with a relatively small number of primordial types, she caused to be born successively, and at diverse epochs, all the animal and vegetable species which inhabit the globe.
>
> . . . .
>
> Nature has operated on an immense scale and with immense resources; we, on the contrary are concerned only with extremely limited means; but between her procedure and ours, between her results and those which we obtain, the difference is all of *quantity;* between her species and those which we create there is only a case of more or less.

We come at last to the famous session of the Linnean Society held on the evening of July 1, 1858, when the Darwin-Wallace theory of natural selection was announced. For twenty years Darwin had been amassing facts to substantiate the theory of evolution and the event which his friends had warned him against had finally occurred, his explanation of evolution had been derived independently by another naturalist, who happened to be his friend, Alfred Russel Wallace. In the spring of 1858 Wallace was at Ternate in the North Moluccas, when during a fever, the full grown conception of natural selection suddenly dawned upon him. He, like Darwin, had been led to the idea by reading Malthus' *Essay*. In a short time he wrote out a description of natural selection in a paper, *On the tendency of varieties to depart indefinitely from the original type*, and sent the paper to Darwin. What then occurred is too well known to need repeating. Sir Charles Lyell and Mr. J. D. Hooker took charge of the situation and insisted that a joint presentation of the theory of natural selection be made by Darwin and Wallace. Darwin's contribution was a short extract from his *Essay* of 1844 together with a letter he had written to Asa Gray in 1857. Wallace's was the paper he had sent Darwin. The proceedings of this session were printed in the *Journ. Linn. Soc.*, **3**: 62, 1859 (Facsimilies of the Darwin-Wallace papers have appeared in *Isis* 14: 133–154, and natural selection had at last received such publicity that no scientist could be ignorant of its implications any longer.

The publication of Darwin's *Origin of Species* late in 1859 definitely ended one era of biology and introduced another. In the meantime, natural selection had been accepted publicly as an explanation of evolution by Joseph Dalton Hooker in his *On the Flora of Australia*, London, 1859.

SUMMARY

As an explanation of evolution, natural selection involves a number of distinct though subordinate propositions, such as the existence of heritable variations, of population pressure, of a struggle for existence with the consequent survival of the fit or the better adapted. Long before Darwin published the *Origin of Species*, a number of philosophers and naturalists had recognized the validity of one or more of these propositions without, however, gaining any clear conception of the implications of natural selection. One such component of natural selection, population pressure, was described by Hale (1677), Buffon (1751), Benj. Franklin (1751), Bonnet (1764), Monboddo (1773), Herder (1784), Smellie (1790), Malthus (1798), Prichard (1808), Wells (1813), Matthews (1831), De Candolle (1833), Lyell (1833), Geoffrey St.

Hilaire (1833), and Spencer (1852). The struggle for existence was described by al-Jahiz (9th cent.), Hobbes (1651), Hale (1677), Buffon (1751), Monboddo (1773), Kant (1775), Herder (1784), Smellie (1790), Erasmus Darwin (1794), Wells (1813), De Candolle (1832), Lyell (1833), and Spencer (1852).

The conception of natural selection itself is very old. It can be used to explain both the evolution of adaptation and the evolution of new species. In the former capacity it serves as an alternative explanation to those facts which are cited as evidences of teleology. Natural selection was described as the cause of adaptation by Empedocles (c. 400 B.C.), Lucretius (99–55 B.C.), Diderot (1749), Maupertius (1756), and Geoffrey St. Hilaire (1833). This explanation was rejected, however, by Aristotle (384–321 B.C.), Lactantius (260–340 A.D.), St. Albertus Magnus (1236), and Whewell (1833). Natural selection was seen as causing the evolution of new species by Wells (1813), Matthews (1831), Darwin (1858), and Wallace (1858). The writings of several savants show that they almost had a clear picture of natural selection but just failed to grasp its full significance. Among these were Rousseau (1749), Prichard (1808, 1826), Lawrence (1819), Geoffrey St. Hilaire (1833), Herbert (1837), Spencer (1852) and Naudin (1852).

A. D. Bache

COMMEMORATION
OF THE LIFE AND WORK OF
ALEXANDER DALLAS BACHE

AND

SYMPOSIUM ON
GEOMAGNETISM

With the cooperation of

The American Philosophical Society
United States Coast and Geodetic Survey
Department of Terrestrial Magnetism of
Carnegie Institution of Washington
Girard College

PHILADELPHIA
THE AMERICAN PHILOSOPHICAL SOCIETY
Independence Square
1941

COMMEMORATION OF THE LIFE AND WORK
OF
ALEXANDER DALLAS BACHE

Alexander Dallas Bache and His Connection with the American Philosophical Society 125

Edwin G. Conklin, Executive Vice-president, American Philosophical Society.

Alexander Dallas Bache and His Connection with The Franklin Institute of the State of Pennsylvania 145

Henry Butler Allen, Secretary and Director, The Franklin Institute of the State of Pennsylvania.

The Connection of Alexander Dallas Bache with the University of Pennsylvania 151

Edward P. Cheyney, Professor Emeritus of European History, University of Pennsylvania.

Bache as an Educator 161

Merle M. Odgers, President, Girard College.

Alexander Dallas Bache as Superintendent of United States Coast Survey, 1843–1867 173

Rear Admiral Leo Otis Colbert, Director, United States Coast and Geodetic Survey.

Alexander Dallas Bache, a Founder, First President and Benefactor of the National Academy of Sciences 181

Frank B. Jewett, President, National Academy of Sciences.

SYMPOSIUM ON GEOMAGNETISM

Terrestrial Electricity in Relation to Geomagnetism 187

O. H. Gish, Assistant Director, Department of Terrestrial Magnetism, Carnegie Institution of Washington.

The Magnetic Survey of the United States 205

N. H. Heck, Chief, Division of Geomagnetism and Seismology, United States Coast and Geodetic Survey.

The Significance of Fossil Magnetism 225

A. G. McNish, Department of Terrestrial Magnetism, Carnegie Institution of Washington.

Geomagnetic Observatories and Instruments 239

H. E. McComb, Chief, Section of Observatories and Equipment, United States Coast and Geodetic Survey.

Magnetic Work at Sea 257

H. F. Johnston, Chief, Section of Observatory Work, Department of Terrestrial Magnetism, Carnegie Institution of Washington.

Geomagnetism: World-Wide and Cosmic Aspects with Especial Reference to Early Research in America 263

J. A. Fleming, Director, Department of Terrestrial Magnetism, Carnegie Institution of Washington.

Aurora and Geomagnetism 299

C. W. Gartlein, Curator, Department of Physics, Cornell University.

Contributions of Ionospheric Research to Geomagnetism 309

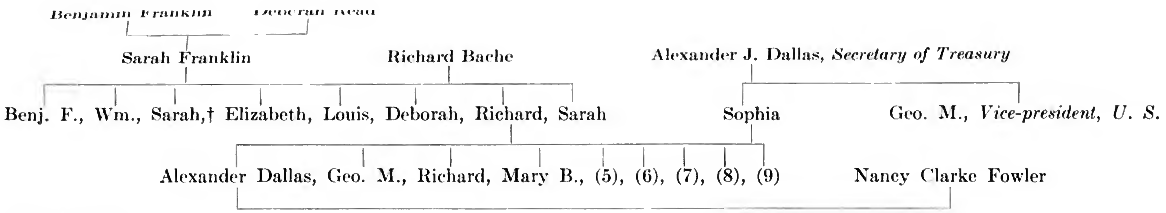
L. V. Berkner, Department of Terrestrial Magnetism, Carnegie Institution of Washington.

Correlations of Short Wave Radio Transmission Across the Atlantic with Magnetic Conditions 323

H. E. Hallborg, Research Engineer, R. C. A. Communications, Inc.

Magnetism and Its Uses 339

Paul R. Heyl, Physicist, National Bureau of Standards.


Benjamin Franklin
Deborah Read
Sarah Franklin
Richard Bache
Alexander J. Dallas, *Secretary of Treasury*
Benj. F., Wm., Sarah,† Elizabeth, Louis, Deborah, Richard, Sarah
Sophia
Geo. M., *Vice-president, U. S.*
Alexander Dallas, Geo. M., Richard, Mary B., (5), (6), (7), (8), (9)
Nancy Clarke Fowler

† Died in infancy

CAREER OF ALEXANDER DALLAS BACHE

1806: Born at Philadelphia, July 19
Student at Clermont Academy, John Sanderson, Preceptor
1821–1825: Cadet in West Point, graduated at 19 as Lieutenant of Engineers
1825–1826: Assistant Professor of Engineering, West Point
1826–1828: Engineering Service at Fort Adams, Newport, R. I.
1828: Married Miss Nancy Clarke Fowler, of Newport, R. I.
1828–1836: Professor of Natural Philosophy and Chemistry, University of Pennsylvania
1829: Elected member, American Philosophical Society
1830: Chairman, monthly meeting, The Franklin Institute
First magnetic observations made in the garden of his home
1832–1857: Secretary, 1832–44; Vice-president, 1845–54; President, 1855–57, American Philosophical Society
1836: Elected President, Girard College
1836–1838: In Europe, investigating educational institutions for Trustees of Girard College
1839: Submitted "Report on Education in Europe"
1839–1842: President, Central High School, and Superintendent of Schools, Philadelphia
1840–1845: Observations in the Magnetic Observatory, Girard College
1842–1843: Again Professor of Natural Philosophy and Chemistry, University of Pennsylvania
1843: Appointed Superintendent of the Office of Weights and Measures of the United States
1843–1867: Superintendent, United States Coast Survey
1846–1867: Regent, Smithsonian Institution
1850–1851: President, American Association for the Advancement of Science
1861–1865: War service, Army and Navy
1863: Planned lines of defense for Philadelphia and supervised construction of some of the works
1863–1867: Founder and President, National Academy of Sciences
1867: Died at Newport, R. I., February 17; his body lay in state in the Hall of the American Philosophical Society. Buried in the Congressional Cemetery, Washington, D. C.

ALEXANDER DALLAS BACHE AND HIS CONNECTION WITH THE AMERICAN PHILOSOPHICAL SOCIETY

EDWIN G. CONKLIN

Executive Vice-president, American Philosophical Society

(Read February 14, 1941, in Commemoration of A. D. Bache)

A FEW words of explanation should be given of the program which has been arranged for this meeting. A year or more ago Captain N. H. Heck, Chief, Division of Geomagnetism and Seismology, United States Coast and Geodetic Survey, suggested that a meeting in commemoration of Alexander Dallas Bache and of the important work on geomagnetism which he started be held at the American Philosophical Society. Our Committee on Meetings gladly approved this proposal, and I take this opportunity of thanking Captain Heck and Dr. John A. Fleming, Director, Department of Terrestrial Magnetism of the Carnegie Institution of Washington, for the important assistance which they have rendered in organizing this program. We are also deeply indebted to Dr. Merle M. Odgers, President, Girard College and to Dr. Roland S. Morris, President, American Philosophical Society and member of the Board of City Trusts in charge of Girard College, for their cooperation in this celebration. It was on the grounds of Girard College, of which Alexander Dallas Bache was the first President, that there was established one hundred years ago under his direction and with aid from the American Philosophical Society the first important magnetic observatory in this hemisphere.

Bache was a leader in so many organizations that it was not possible to have all of them represented individually on this program. Six organizations with which he was intimately associated are represented by officers of those bodies on this program, namely, the American Philosophical Society, the University of Pennsylvania, the Franklin Institute, Girard College, the United States Coast and Geodetic Survey and the National Academy of Sciences. A sixth institution which is represented by the speaker on the program this evening, namely, the Department of

Terrestrial Magnetism of the Carnegie Institution of Washington was not in existence in Professor Bache's time, but it represents the splendid development of work which he started. A seventh institution represented by the presiding officer tomorrow morning and by one of the speakers tomorrow afternoon is the National Bureau of Standards, which is an outgrowth of the Office of Weights and Measures of which Bache was Superintendent.[1]

Other organizations that should have been represented on the program if time and circumstances had permitted are the United States Military Academy at West Point, the Central High School of Philadelphia, the American Association for the Advancement of Science and the Smithsonian Institution; most of these are represented by delegates. Professor Bache was closely associated with all of these organizations and he was the inspiring leader of some of them.

Benjamin Apthorp Gould said of him in his eulogy before the American Association for the Advancement of Science (*Proc. A. A. A. S.*, 1868):

> He stood forth preeminent as our leader in science, our first counsellor where her welfare was at stake . . . I know of no department of physical or natural science which has not been stimulated or fostered by his means. The legislative and executive departments of the government, the army and navy, the progress of education and scientific discovery, the growth of the arts, the spread of commerce—all have been greater in America because he has lived.

After this general introduction I shall as far as possible limit my remarks to the specific topic assigned me, namely, Bache's connection with the American Philosophical Society, in order to avoid covering the ground assigned to other speakers.

The American Philosophical Society is proud to take a leading part in this commemoration of one of America's great scientists and one of its most distinguished members. He was elected to membership on April 17, 1829, when he was not yet twenty-three years old. He attended the next meeting on May 1, and began at once to take an active part in the meetings and committees. For twenty-five years he was an officer of the Society, being

---

[1] A collection of Bache's instruments, models, manuscripts, and books furnished by the American Philosophical Society, the United States Coast and Geodetic Survey and the Department of Terrestrial Magnetism of the Carnegie Institution of Washington was exhibited in the Hall of the American Philosophical Society in connection with this meeting.

Secretary from 1832 to 1844, Vice-president, 1845–1854 and President, 1855–1857.

He belonged to this Society by inheritance as well as by election, for Benjamin Franklin was his paternal great grandfather; Richard Bache, Postmaster General of the Colonies (1776–82) and Franklin's son-in-law, was his grandfather. An uncle, William Bache, M.D., was a member, and a cousin, Franklin Bache, M.D., Professor of Chemistry in Jefferson Medical College, was a member, Vice-president and President of the Society only a few years before Alexander Dallas Bache occupied these offices and the two were always distinguished by referring to the former as "Dr." and the latter as "Prof." Bache. On his mother's side both his grandfather, Alexander James Dallas, Secretary of the Treasury in Madison's administration, and his mother's brother, George M. Dallas, Vice-president of the United States and Minister to Great Britain, were long time members of this Society.

From the date of his election to this Society in 1829 until his departure for Europe in 1836 there were few meetings of the Society when his name does not occur in the Minutes as taking some active part in the meeting, and after his return from Europe in 1838 until his departure for Washington in 1843 he presented papers, verbal communications, letters from foreign correspondents and observatories, proposals from the Royal Society of London and the British Association for the Advancement of Science in relation to Magnetic Observatories in the United States, at many meetings. Some of the most interesting records in the Minutes on these subjects are the following:

> September 20, 1839. Professor Bache made a verbal communication of the measures taken by the British government, on the recommendation of the British Association, and under the advice of the Royal Society, for obtaining a series of magnetic observations in different quarters of the globe, in conjunction with a naval expedition in the southern hemisphere, under the command of Capt. James Clark Ross, and read extracts from letters of Professor Lloyd and Major Sabine, relating to the preparation for the undertaking.
>
> Professor Bache further stated, that on submitting the circular addressed to him by the Foreign Secretary of the Royal Society, with extracts from the letters before referred to, and other information as to the nature and importance of the results to be obtained by this combined system of magnetic observations, to the Building Committee of the Girard College, through their Architect, they had, with creditable liberality, given orders for the erection of an observatory suited to the obser-

vations contemplated, and to the instruments already in the possession of the Trustees of the College.

Professor Bache submitted the plans of the observatory, drawn by Thos. U. Walter, Esq. Architect. (*Amer. Philos. Soc. Proc.*, I: 117–118.)

November 15, 1839. The following resolutions in relation to combined magnetic observations were adopted:

Resolved, That in the opinion of the American Philosophical Society, it is highly desirable that the combined series of magnetic observations now in progress under the direction of the British government, should be extended to the United States, by the establishment of Magnetic Observatories at suitable places.

Resolved, That a Committee be appointed, with authority, on behalf of the Society, to invite the attention of one of the departments of the Government of the United States to the plan for combined magnetic observations, a sketch of which was presented in the documents from a Committee of the Royal Society of London, and to urge cooperation in the plan as a national undertaking, in every way worthy of the United States.

The Committee under the above resolution, consists of Professor Bache, Doctor Patterson, Professor Henry, Mr. Kane, and Colonel Totten. (*Amer. Philos. Soc. Proc.*, I: 148.)

December 6, 1839

Professor Bache also presented, on behalf of the Committee on Magnetic Observations, appointed at the last meeting of the Society, a Memorial, addressed to the Secretary of War of the United States, inviting the establishment of Magnetic Observatories, and cooperation in the system of magnetic and meteorological observations now in progress under the direction of the British government. (*Amer. Philos. Soc. Proc.*, I: 151.)

January 17, 1840

Professor Bache read extracts from a letter from Major Sabine, of England, describing the progress of measures for contemporaneous magnetic and meteorological observations for the next three years, urging a concert of observations in the United States, and stating that a magnetic survey of the British dominions north of the United States would be made, and suggesting a similar survey of the United States. Professor Bache remarked, that steps had been taken some months since, by several gentlemen who have heretofore been engaged in magnetic researches, to procure such a survey, the result of which, however, was yet uncertain. (*Amer. Philos. Soc. Proc.*, I: 169.)

March 20, 1840

The Committee, consisting of Professor Henry, Dr. Patterson, and Mr. Walker, to whom was referred a paper entitled, "Observations of the Magnetic Intensity at twenty-one Stations in Europe. By A. D.

Bache, LL.D. President of the Girard College for Orphans, &c.," reported in favour of the publication of the paper in the Society's TRANSACTIONS. The Report was adopted, and the publication ordered accordingly.

The stations at which the observations, recorded in this memoir, were made, were twenty-one in number: three in Great Britain, and the others in the continent of Europe. They include Edinburgh, Dublin, London, Brussels, Berlin, Paris, Vienna, the Flégière, Brientz, the Faulhorn, Geneva, Chamberi, Chamouni, Lyons, Milan, Venice, Trieste, Florence, Turin, Rome and Naples. . . . [There follows a summary of these observations and a table giving the horizontal intensity and dip for each station.] (*Amer. Philos. Soc. Proc.*, I: 185.)

July 3, 1840

Professor Bache, from the Committee on magnetic observations, read an extract from a letter of Major Sabine, V. P. of the Royal Society of London, stating that the Council of the Society had, on the recommendation of the Committee of Physics, expressed their opinion of the importance to the plan of combined magnetic observations now in progress, that observatories should be established in the United States, and had instructed their President to bring this expression of opinion to the knowledge of the government of this country.

Professor Bache stated that the resolution just referred to had been adopted with a view to aid the efforts of this Society in procuring the erection of observatories, as recommended in their Memorial to the Secretary of War, which had been referred, by that officer, to Congress.

He also read an extract from a subsequent letter from Major Sabine, in reference to the progress of the combined magnetic observations, stating that the Emperor of Russia had ordered the erection of nine magnetic and meteorological observatories in his dominions, to conform, in respect to instruments and times of observations to the system recommended by the Royal Society. One of these observatories is to be upon the N. W. coast of America [Alaska].

Professor Bache stated that the regular system of bi-hourly magnetic and meteorological observations was now established in the observatory at the Girard College, and had been in progress since the close of the month of May. He intended, at a future day, to present to the Society the names of the gentlemen, chiefly members of the American Philosophical Society, by whose contributions a fund had been raised to defray the expense of employing the assistants required for these observations. . . . (*Amer. Philos. Soc. Proc.*, I: 242.) [See Minute of May 21, 1841.]

November 6, 1840

Professor Bache read an extract of a letter from Lieut. Riddell, Director of the Magnetic Observatory at Toronto, U[pper] C[anada], which stated that an entire discordance had been found between the curve representing the changes of inclination, on the February magnetic term day, at Toronto, Dublin, Brussels, and Prague, whilst those at the last three named stations agreed very well together. This result, Professor B[ache] stated, confirms the conclusions previously drawn from the observations at short intervals, of Prof. Lloyd and himself, in November last.

Mr. Walker made some observations in relation to the Observatory of the Harvard University, Cambridge, and stated that extensive arrangements had been made, and were in contemplation, for prosecuting magnetic observations and practical astronomy.

Professor Bache made a verbal communication of some recent determinations of the magnetic dip, made by him at Philadelphia and Baltimore. (*Amer. Philos. Soc. Proc.*, I: 293–294.)

December 4, 1840

Professor Bache brought before the Society an instrument for measuring the changes in the vertical components of the force of terrestrial magnetism, which he described as combining the principles of the vertical force instrument of Prof. Lloyd, with that of reflection adopted in the magnetometers of Prof. Gauss, and which had been made for him by Mr. Saxton.

Professor Bache called the attention of the Society to a diagram representing the changes of magnetic declination, as recorded at the Magnetic Observatory of Mr. Bond, at Cambridge, and at the Girard College, on the magnetic term day of May, 1840, and showing that the changes attending the aurora are not peculiar to one locality, but that, as observed at different places, they are parts of a great magnetic disturbance.

Professor Bache read a letter from Major Sabine, giving the progress of the magnetic observations now making, and referring to the modes deemed advisable for the publication of the records of observatories. He referred also to the anomalous nature of the curves for the May term day at Toronto and at Greenwich, and to an instrument for observing vertical force by reflection, in the putting up of which Professor Airy was engaged. (*Amer. Philos. Soc. Proc.*, I: 310–313.)

December 18, 1840

Professor Bache laid before the Society a Report from Mr. [J. Q.] Adams to the House of Representatives, on a Letter from the Secretary of War, of the 31st Dec. 1839; and a Memorial from a Committee of the American Philosophical Society, asking the aid of the government to carry on a series of magnetic and meteorological observations, and ending with a resolution:

"That the sum of twenty thousand dollars ought to be appropriated for the establishment of five several stations, at suitable distances from each other, for making observations of terrestrial magnetism and meteorology, conformably to the invitation from the Royal Society of Great

Britain to the American Philosophical Society at Philadelphia, and to other learned societies in the United States; that the said sum should be placed under the direction, and at the disposal, of the Secretary of War, for the fulfilment of these purposes; he to account for the expenditures, thus authorized, to the Treasury of the United States."

Professor Bache then offered the following resolution, which was adopted:

Resolved, That the Committee by whom a memorial was addressed to the Secretary of War, in reference to the establishment of magnetic observations, be instructed again to call his attention to the system of combined observations on terrestrial magnetism and meteorology now in progress. (*Amer. Philos. Soc. Proc.*, I: 320.)

January 15, 1841

Professor Bache read an extract from a letter of Major Sabine, stating that the changes of magnetic declination, and of horizontal force, would be observed on the term days, with transportable magnetometers furnished by the British Association, by Mr. Schomberg, at Demarara, in Guiana; this being the first magnetic station yet established in South America.

Professor Bache also stated, that he had received from Mr. Bond, abstracts of the term day observations of changes of magnetic declination at Cambridge, Mass., for the months of June, July, August and September; also, that he had received from Lieut. Gilliss, of the U. S. Navy, observations of declination made at Washington, from the 5th to the 9th of January, inclusive, from 9 to 10 A.M. of Göttingen time at short intervals, for comparison with similar observations at Philadelphia and Toronto. These observations he proposes, when compared with those at the Philadelphia Magnetic Observatory, to communicate to the Society. (*Amer. Philos. Soc. Proc.*, II: 6.)

February 5, 1841

Professor Bache presented to the notice of the Society observations of magnetic declination received from Mr. Bond, which had been made by that gentleman during six days, commencing on the 4th of January, in concert with the Magnetic Observatories at Philadelphia, Washington and Toronto. (*Amer. Philos. Soc. Proc.*, II: 15–16.)

May 21, 1841

Professor Bache communicated to the Society a statement of the Observations made for the year past at the Magnetic Observatory at the Girard College, and exhibited the original records, the abstracts made from them, the calculated results, and the curves by which they are represented. He reminded the members that in consequence of the depressed state of the Society's funds in May last, it had been judged inexpedient to ask for the appropriation of any part of them to the object of these observations; and he mentioned the names of ten members of the Society, and of three gentlemen, not members, Messrs. Richard Price and J. D. Brown, of Philadelphia, and Professor M'Lean, of Princeton, by whose liberality the Observatory had been supported during the year.

After some remarks from Mr. Walker, describing the results which have been arrived at by the labours of Gauss, Weber and others, in magnetism, and referring to the practical value to navigation of the magnetic investigations now making, Dr. Chapman pressed upon the Society the importance of continuing the magnetic and meteorological observations in the combined series which is now in the course of execution: . . . and on his motion, a committee was appointed to devise means for continuing the observations at the Girard College Observatory during the remaining two years of the series. (*Amer. Philos. Soc. Proc.*, II: 69.)

June 25, 1841

[Special meeting for consideration of report of Committee to devise means for completing magnetic and meteorological observations.]

In this report the Committee review the history of the concerted observations first proposed by the Royal Society of London in 1839, and refer to the action of this Society in consequence of the circular from that learned body. They express the strongest confidence in the skill, assiduity, and success, with which the operations of the magnetic observatory at Girard College have been thus far conducted, and a belief that their prosecution is called for by the honour of the Society. They pledge themselves, that as the funds required for the past year's expenditure have been furnished by individual contributions of the members and their friends, so those for the third year shall be supplied in like manner, if the Society will defray the charges of the intervening period from its corporate funds.

After a full discussion of the subject, resolutions were adopted by the Society—

1. Directing the Committee on the Observatory to ask permission from the City Councils of Philadelphia to constitute, from certain moneys heretofore paid by the city, a fund for the promotion of astronomical and magnetic researches, and the publication of the results thereof.
2. Authorizing the Committee to refund, if required, certain contributions made by individuals to the Astronomical Fund; and
3. Directing the trustees of that fund to supply the means of completing the magnetic and meteorological observations on certain terms and conditions. (*Amer. Philos. Soc. Proc.*, II: 77–78.)

January 21, 1842

Professor A. D. Bache communicated, on behalf of Mr. Nicollet, of Washington, an abstract of observations on the magnetic dip, made at Baltimore, Washington, Philadelphia, Albany, Oswego, Niagara Falls, Detroit, Mackinaw Island, Chicago, Joliet, Ottawa, Peru, Illinoistown and St. Louis. He also read a letter from Major Sabine, communicating the progress of the general series of magnetic observations; and one from Prof. Loomis, of Western Reserve College, stating that, last autumn, he had made observations of the magnetic dip at nearly forty different stations, in the north-west part of the United States, the results of which he intended hereafter to communicate to the Society. (*Amer. Philos. Soc. Proc.*, II: 144.) [See Loomis, *Amer. Philos. Soc. Trans.*, VII: 1–6, 101–111; VIII: 61–72, 285–304; IX: 181–186].

March 4, 1842

Professor A. D. Bache exhibited the curves representing the results of the bi-hourly magnetic observations, made during the years 1840 and 1841 at the Girard College Observatory, showing the daily changes of magnetic declination, and horizontal and vertical intensity. He stated that from these curves the approximate times of maximum and minimum could be inferred; but that, in order to render the determination of the periods of their occurrence more accurate, additional observations at every six minutes were now made (since January 1) within the limits shown by the curves, presented this evening, to be those of the occurrence of maxima and minima. (*Amer. Philos. Soc. Proc.*, II: 150–151.)

October 7, 1842

Professor Bache drew the attention of the Society to the necessity of providing means for continuing the observations now making under the direction of the Society at the Magnetic Observatory, or of closing the Observatory: whereupon, on motion of Dr. Patterson, a special Committee was appointed, to report in regard to providing means for the continuation of the observations. Committee, Dr. Chapman, Dr. Patterson, Dr. Wood, Mr. Fraley, and Mr. Kane. (*Amer. Philos. Soc. Proc.*, II: 222.)

November 18, 1842

Professor Bache stated that his attention had recently been particularly called, by a letter from M. Quetelet, Secretary of the Academy of Sciences of Brussels, to the general instructions for simultaneous observations of natural phenomena, issued by the Academy. Co-operation in the system of observations by observers in the United States being very desirable, Prof. Bache asked leave to offer the following resolution:

"*Resolved*, That a Committee of five members be appointed, to report to the Society what measures may be taken, most effectually to secure co-operation, by observers in the United States, in the system of observations of periodical natural phenomena, forming the subject of the instructions of the Brussels Academy of Sciences." (*Amer. Philos. Soc. Proc.*, II: 235.)

This resolution was adopted, and the following Committee appointed: Prof. Bache, Dr. Patterson, Mr. Frazer, Dr. Griscom and Mr. Lea.

January 6, 1843

Professor Bache communicated to the Society, that in consequence of the want of funds for the support of the Magnetic Observatory, the bi-hourly observations, and those for maxima and minima, had ceased with the first of the present year.[1] He further stated, that it was his intention to keep up the term-day observations, and also to have an observation of the magnetic instruments made each day, to connect the indications of the magnetometers from one term-day to another. Certain of the meteorological observations were also to be continued. (*Amer. Philos. Soc. Proc.*, II: 247.)

April 7, 1843

Professor Bache communicated an extract of a letter from M. Quetelet of Brussels, stating that hourly meteorological observations were made at some fifty stations in Europe at the periods of the equinoxes and solstices, the observers corresponding with the Academy of Sciences of Brussels. M. Quetelet expresses the wish that the American Philosophical Society should become the centre of a similar union for the new world, and urges that the attention of men of science should be called to the subject; he also enforces the necessity for conformity to the plan laid down in the circular of the Academy. (*Amer. Philos. Soc. Proc.*, II: 266–267.)

Professor Bache informed the Society that he had received a letter from the Hon. James M. Porter, Secretary of War, through the bureau of Topographical Engineers, making an allowance for the continuance of the observations at the magnetic observatory. In consequence of this liberal and well-timed supply of means, he said, the series of observations was resumed on the first of the present month.

The Society thereupon unanimously adopted a resolution, tendering its thanks to the Hon. James M. Porter, Secretary of War, for his judicious appropriation of means for the continued prosecution of the magnetic observations heretofore conducted under the auspices of the Society: and

The Committee of the Magnetic Observatory was instructed to communicate the resolution to the Hon. Secretary. (*Amer. Philos. Soc. Proc.*, II: 268.)

May 27, 1843

Professor Bache presented the results of two years' observations of the magnetic elements, and of the temperature, pressure, and moisture of the atmosphere, at the Magnetic Observatory at the Girard College.

The Observatory was opened in May, 1840, and the results are for two years from that time. Professor Bache, after alluding generally to the different results obtained from the magnetic observations, called particular attention to the averages of the hourly changes. The observed elements were the declination, and the horizontal and vertical forces. The changes of each of these were represented by curves in the usual way, the differences of the abscissae representing the times elapsed, and the differences of the ordinates the change of element; increasing ordinates corresponding to increase of declination or of force. The curves, even with the limited number of observations upon which they are founded, are remarkable for their regularity.

---

[1] Subsequently resumed.—Reporter.

The declination shows distinctly two maxima and two minima during the twenty-four hours. The chief maximum is about 1 P.M., and the subordinate maximum about 1 A.M.; the latter, though marked, does not reach the mean line of the twenty-four hours. The minima are at about 8 A.M. and 9 P.M. . . .

Professor Bache gave reasons why he considered the phenomena thus presented to be inconsistent with the idea, that they are produced by the heating effects of the sun, acting either directly or indirectly. . . . Professor Bache observed, in conclusion, that as the Observatory had been re-opened by aid obtained from the Secretary of War, similar observations would by degrees be accumulated, and by an increased number of results render the means more trustworthy than those now presented.

At the close of his remarks, Professor B[ache] invited the members and correspondents of the Society and the gentlemen attending its meetings, to inspect the Observatory at the Girard College after the adjournment in the afternoon. (*Amer. Philos. Soc. Proc.*, III: 90–92.)

May 2, 1845

Professor A. D. Bache called attention to the report made by him in February last, to the Treasury Department, on the progress of the construction of standard weights and measures, and of balances. A copy of this report had been presented to the Society at a previous meeting.

The work of constructing standards had been commenced by the late Mr. F. R. Hassler in 1835; and at the time of his decease the standard weights for the Custom Houses of the United States and for the States had been made, and generally delivered. One-third of the capacity measures had been completed, and the rest were in different stages of progress. About one-fourth of the measures of length had been finished, and the rest were in progress. Many other standards for miscellaneous purposes had been made and delivered. The balances used in the office of weights and measures to adjust standards has been made; two other balances had been finished, and thirty, intended for distribution to the States, had been commenced. . . .

Much progress having been made in the preparation and distribution of standard measures when Professor Bache came into the charge of the work, he had deemed it necessary to adhere to the methods of Mr. Hassler, and to use the tables founded upon his experiments, for reductions; otherwise the uniformity of the system would have been destroyed. . . . (*Amer. Philos. Soc. Proc.*, IV: 159–160.)

Throughout his whole life Bache showed an amazing activity in teaching, experimenting, writing, administration. Here in Philadelphia his time was divided among the University of Pennsylvania, the American Philosophical Society, the Franklin Institute, Girard College, the Observatory and the High School. Everything he did was excellently done. He was an inspiring and original teacher, a careful and exact observer and experimenter, a clear and direct writer and a farseeing and genial administrator. How he ever found time to do all of this important work in a "horse and buggy era" is a constant surprise to us of the flying age.

There are one hundred and forty-seven publications listed in his bibliography; twenty-three of these appeared in the *Journal* of the Franklin Institute between 1829 and 1836; twenty-one appeared in the TRANSACTIONS and PROCEEDINGS of the American Philosophical Society after 1836; nine in the *American Journal of Science;* thirty-six in the *Proceedings* of the American Association for the Advancement of Science; eight in the Smithsonian Contributions to Knowledge; thirty-one in the Reports of the United States Coast Survey. Others will speak of his scientific work in connection with his various official positions and I shall not undertake to report upon even that portion which was published by this Society, but I may properly refer to the fine quality of the first monograph which he read before the Society when he was only twenty-six years old and which was published in our TRANSACTIONS, V: 1–21. It shows a breadth of view, a maturity of judgment and a becoming modesty such as are not always manifested by brilliant youth. In this same volume of the TRANSACTIONS he published three other articles on magnetism and meteorology.

The extracts from the Minutes of the American Philosophical Society, which have been given, indicate briefly the part which this Society took in the support of his Magnetic Observatory and in the international work in geomagnetism. The results of that work were reported by Professor Bache in many minor papers in the PROCEEDINGS and in a few major communications in the TRANSACTIONS (see especially Vol. VII: 75–100), most of which appeared before he removed to Washington to become Superintendent of the United States Coast Survey.

In his address before the American Association for the Advancement of Science (1868) in Commemoration of Bache, Benjamin A.

Gould speaks of the numerous experiments which he had made at his own home in Philadelphia, (which in our TRANSACTIONS, I find was at "Chestnut Street near Schulkill Sixth Street", at present 17th Street). Among these were a series of experiments with Fourier's thermoscope of contact. But finding these results unsatisfactory, he had invented a new thermoscope of contact better adapted to the purpose. He carried his investigations on heat to a considerable extent, but never completed them. Gould relates the following incident to make known the reason of their discontinuance:

> One room on the sunny side of his house was appropriated to these experiments; the various thermoscopes and all the subsidiary apparatus were arranged there, and the apartment was held sacred to scientific investigation. One evening, while he was attending a session of the Philosophical Society, an alarm of fire broke out in the neighborhood. His mother, then a member of his family, heard the alarm, and hastily entered the room without a lamp, to look from the front window. A crash reminded her, too late, of the inconsiderateness of her movements. The apparatus was entirely destroyed, and the first words which greeted her son on his return told him what he had lost. He made no reply, but went to the room and silently surveyed it. The destruction was complete, and the hard labors of nearly a year were rendered fruitless. An eyewitness has described it to me. He stood white with emotion for a few moments; then, turning away, only trusted himself to say that he would return soon, and hurried out of the house. Half an hour in open air restored him to himself; returning, he consoled his mother, and made light of the occurrence; nor did he ever afterwards explain the reason why his observations on heat were discontinued.

In 1807, President Jefferson had recommended to Congress the establishment of a Coast Survey, and it is permissible to think that he had been urged to this action by his fellow member in the American Philosophical Society, Ferdinand R. Hassler, Swiss Engineer, who had recently arrived in this country. The American Philosophical Society had approved Hassler's plans, but no action was taken by Congress until 1817 when the Coast Survey was established and Hassler appointed Superintendent.

On March 3, 1820, Hassler communicated to the Society a large monograph entitled "Papers on Various Subjects Connected with the Survey of the Coast of the United States" (TRANS., II: 232–418). He began with a letter from Albert Gallatin, Secretary of the Treasury and for sixty years a member of this Society, dated "Treasury Department, 25th March, 1807," and saying, "The President of the United States being authorized by an Act of the last Session to cause the whole of the coast of the said States, together with the adjacent shoals and soundings to be surveyed—has directed me to apply to you, requesting that you would have the goodness to suggest the outlines of such a plan," etc.—

To this letter Hassler replied in a letter in French, dated "Philadelphia, 2 Avril, 1807" which occupies pages 234–240 of Trans., Vol. II. This was followed (nine years later, Jan. 5, 1816) by Hassler's "Plan for putting into operation the Survey of the Coast of the United States" and a "Catalogue of the Instruments and Books collected for the Survey of the Coast," a "Comparison of the French and English Standard Measures" etc., "Description of the Apparatus for measuring Base Lines," "Description of the Two-foot Theodolite," and a long account of the use of that instrument, telescope, clocks, etc., "Plan of an Observatory proposed to be built in Washington", "Mechanical Organization of a Large Survey" and nine pages of "Examplars of the Day Book and Journal of Results" with eight full-page copper-plate engravings by Caroline Hassler of instruments described in the article—a total of nearly two hundred pages. At the end of this long and painfully detailed article the following note was added by the Editors of the Transactions, p. 419:

> This Paper was followed by a Journal of the "Principal Dates connected with the Survey of the Coast," but as this Journal was not considered of general interest, and as the paper was already of great length, it has been thought proper to omit it.

All this indicates how great was the interest of this Society in the establishment of the Coast Survey.

In his eulogy of Alexander Dallas Bache, Joseph Henry says that when Hassler died in 1843 little more than a beginning had been made in the Survey and that the work was liable to be suspended or abolished by Congress at any time. He adds,

> His [Bache's] appointment to this position was first suggested by the members of the American Philosophical Society, and the nomination fully concurred in by the principal scientific and literary institutions of the country. In this movement he himself took no part, and indeed regarded the position as one not to be coveted; for while it opened a wide field for the exercise of talent and the acquisition of an enviable reputation, it involved responsibilities and presented difficulties of the gravest character. Professor Bache was not one of those who, abounding in

self-confidence, imagine themselves equal to every exigency, or who seek the distinctions and emoluments of office without any regard to the services to be rendered or the duties to be discharged.

The friends of Professor Bache rallied to his support. Professor Henry wrote, "No other living man is so well qualified as himself to fill this office." The following manuscript letter in the Gratz Collection of the Historical Society of Pennsylvania from John Torrey to Benjamin Silliman shows the high regard in which this young professor was held by older scientists:[2]

Manuscript letter from J[ohn] Torrey to Prof. B[enjamin] Silliman
New York. Nov. 29th, 1843.

My dear Sir:

You may have heard that Prof. A. D. Bache is applying for the situation left vacant by the death of Mr. Hassler (Supt. of the Coast Survey). He wishes your recommendation in his behalf & would have written to you himself on the subject had he not thought it would have been asking too much of you. His reputation is so well established in the world of science, that among his peers no testimonials would be required, but not so with the great men at Washington. You know that Bache left the Military Academy in 1825—served under Col. Totten in the Corps of Engineers for some time—was called to the Prof^{r}—of Natural Philosophy in the University of Pennsylvania, then to the Presidency of Girard College—afterwards was appointed to Superintend the High School in Philadelphia—& lastly reappointed in the University. You are familiar with most that he has done—& know that few men in this Country stand higher than he as a nat. philosopher. He is a practical man also, & is perfectly well acquainted with the use of all kinds of scientific instruments. He is a gentleman & a scholar. Now if you can say this much in behalf of Prof. B. it may be of the greatest consequence to him & materially aid in procuring his appointment. Your name would have much weight with the President, & others at Washington. The Secty. of the Treasury will have much to say in the matter & the Secty. of the Navy will also have a voice.

His removal from Philadelphia interfered with, but did not break his close connections with this Society. He retired from the office of Secretary which he had occupied for twelve years (1832–1844), but was at once elected Vice-president, which office he held for nine years (1845–1854) when he became President (1855–1857). He presided at the meetings of the Society from time to time and occasionally presented summaries of the work of the Coast Survey and of the Office of Weights and Measures, but his time and efforts were more and more absorbed by his many activities at the seat of government. This was especially the case during the Civil War when he was almost continually occupied with duties for the Army and Navy, the Sanitary Commission, of which he was Vice-president, and the newly created National Academy of Sciences, of which he was the chief founder and first President.

---

[2] I am indebted to Dr. Alban W. Hoopes for calling my attention to this letter.

However, he always regarded Philadelphia as his real home and when called upon by the Mayor of the City to superintend the plans for the protection of Philadelphia against the invasion of Lee's Army, which culminated at Gettysburg, he spent that sultry summer overseeing this work, and thereby impaired his health which continued to decline until his death on February 17, 1867.

The Minutes of the American Philosophical Society referring to his death and the arrangements for receiving his body in the Hall of the Society are given herewith:

February 19, 1867
(Special Meeting)

The President stated that intelligence of the death of Professor Alexander Dallas Bache, on the 17th inst., had been received, and that he had convened this meeting for the purpose of taking such measures as might be deemed proper for the loss of so distinguished an associate, and to tender the use of the hall to the committee appointed by a meeting of the scientific, learned, and other bodies of the city, held at the Chapel of the University of Pennsylvania this day.

Mr. Fraley then reported to the Society the proceedings of said meeting, and gave a feeling and appropriate sketch of the life, character, and services of Professor Bache, and offered the following resolutions, which were adopted:

Resolved, That a committee be appointed to prepare appropriate resolutions, expressive of the regret and sorrow of the Society at the loss it has sustained by the death of Alexander Dallas Bache.

Resolved, That the hall of the Society be placed at the disposal of the committee, this day appointed at a meeting held by the members of the learned Societies and other bodies, for the reception and care of the remains of Professor Bache, while they remain in this city.

Resolved, That the chair of the presiding officer, and the hall of the Society, be draped in mourning for the space of six months.

Resolved, That a committee of three be appointed to co-operate with the committee above named, in arranging for the reception and proper care of the body of Professor Bache while it remains in the hall of the Society, and for making any arrangements for any funeral procession, if such should be deemed proper.

The committee appointed under the first resolution consisted of Mr. Fraley, Mr. Fisher and Mr. Peale.

The committee appointed under the last resolution consisted of Mr. Cuyler, Mr. Price and Mr. T. P. James. (*Amer. Philos. Soc. Proc.*, X: 326–327.)

March 1, 1867

Mr. Fraley . . . reported, on behalf of the committee appointed at the special meeting of the 19th ultimo, the following resolutions, which were adopted:

Resolved, That by the death of Alexander Dallas Bache, the American Philosophical Society has lost a member distinguished for his zeal in, and devotion to, the cause of Science, and for the excellency of his private life and character.

Resolved, That the reputation won by Professor Bache in the cause of education, and in extending the whole realms of science and useful knowledge is a source of honorable gratification to his friends and associates, and is to be most highly prized as a matter of national pride, and as placing our country, in these respects, on an equality with any other in the world.

Resolved, That the termination of the life of such a man is deeply to be mourned, for it rarely happens that so much mental strength and such stores of knowledge are combined with such winning manners, such delicate and profound discrimination in the choice of men and places, and such great executive ability in accomplishing results and awarding to every colaborer his full measure of appreciation and honor.

Resolved, That this Society, connected as it has been with him from his early manhood until his death, and having had the benefit of his services and counsel in an eminent degree during nearly the whole of that period, is especially called upon to deplore his loss, and to perpetuate on its records the memory of one so wise and good, and to keep him constantly in remembrance as an example worthy of general imitation.

Resolved, That a member be appointed to prepare an obituary notice of Professor Bache for publication in the PROCEEDINGS of the Society.

Resolved, That the officers of the Society be requested to forward a copy of these proceedings, and of those of the 19th ultimo, to the widow of Professor Bache, and to express our cordial sympathy with her in the great bereavement with which she has been visited.

On motion of Professor Frazer, Mr. Fraley was appointed to prepare the obituary notice. (*Amer. Philos. Soc. Proc.*, X: 327–328.)

April 19, 1867

Letters were read . . . from a number of gentlemen of Philadelphia, dated April 11th, 1867, asking the co-operation of the Society in the erection of a monument to the memory of Professor A. Dallas Bache.

Resolved, That a Committee be appointed to collect subscriptions for the erection of a monument to the memory of Professor Alexander Dallas Bache.

The committee consists of Prof. Kendall, Mr. Farley and Mr. Thomas P. James. (*Amer. Philos. Soc. Proc.*, X: 335–336.)

[No record of the erection of this monument has been found.]

To these resolutions of the Society which he had served and honored so long may be added the editorial comments of two Philadelphia newspapers, which Dr. Alban W. Hoopes of the Library Staff of this Society has brought to my attention:

Philadelphia Inquirer, Monday, February 18, 1867

Editorial

We deeply regret to announce the death of Alexander Dallas Bache, at Newport, yesterday. He was the great grandson of Benjamin Franklin, and was born in this city. He held the post of Professor in the University, until he was sent to Europe on behalf of the Trustees of Girard College, to visit the institutions of learning there, and report a plan for the organization of that College. On his return he made a very full report on education in Europe, which was published by the city, and which was a most valuable work. During the completion of Girard College he accepted the Principalship of the Central High School, and reorganized that institution. He established the division of studies since pursued there, trebled the departments, and established the system of semi-annual examinations and semi-annual commencements.

After remaining at the head of that school for three years, he was called to the Chair of Natural Philosophy and Chemistry at the University of Pennsylvania. He remained there for two years only, having been appointed by President Tyler Superintendent of the Coast Survey, then vacated by the death of Mr. Has[s]ler. Since that time his name has been identified with that great American work, the coast survey, which, under him, was developed in extent, and in scientific and practical completeness, beyond that of any other country in the world.

The reputation of Professor Bache is world-wide. He has been complimented with honorary membership in the leading scientific societies of Europe, and has, from time to time, received tokens of honor from them. He enjoyed the personal friendship and correspondence of such men as Humboldt, Farraday [sic], Agassiz, Henry, Pierce and Graham; and in this country his name is identified, as an active promoter, with nearly every scientific body. He was one of the presiding officers of the American Philosophical Society, one of the founders of the Franklin Institute, and also of the National Association [Academy] of Science and the Academy of Natural Sciences.

When the Rebellion broke out, he resisted every effort on the part of the Southern leaders to win him over as they had done Maury and others. He baffled their secret attempts to obtain maps and charts of the Southern harbors and coasts from the Coast Survey Office. At the same time he rendered most efficient service in providing our blockading squadrons with accurate charts of the coast and its inlets, and the army with valuable topographical maps, compiled and published under his direction.

About the time of the invasion of this State by the Rebels, he was invited to this city to superintend the fortifications in process of erection. It was while devoting himself to this work that, by severe intellectual labor and exposure, he so impaired his health as to lay the seeds of the disorder from which he died at Newport yesterday.

He was a most thorough man of science, eminently proficient in mathematics, natural philosophy, and chemistry, and of remarkable administrative ability.

Modest and quiet in his demeanor, he was zealous in the promotion of science and education, and in private life he was a most devoted husband and a sincere Christian. He was warmly attached to Philadelphia and her interests, and we are indebted for the selection of League Island for our national navy yard, in a very great degree, to his well directed and persistent efforts while acting as one of the first commissioners to report on a site for a navy yard.

His loss will be severely regretted by a large circle of friends here and by all true men of science throughout the world.

Philadelphia Inquirer, February 22, 1867
Obsequies of the late Professor Bache

Throughout yesterday the body of Professor Alexander Dallas Bache, LL.D., late Superintendent of the Coast Survey, lay in state in the rooms of the Philadelphia Philosophical Society, on Fifth Street, below Chestnut. The halls were tastefully draped in black, and the body placed in the south room. . . .

Owing to the festivities incident to today as the anniversary of Washington's birthday, the remains will be escorted unostentatiously to the Baltimore depot, leaving the hall at ten o'clock A.M.

The Age, Tuesday, February 19, 1867

Editorial

We have to record the death of an eminent Philadelphian, Alexander Dallas Bache, which took place at Newport a few days ago. Mr. Bache was a Philadelphian by birth, family, boyish education, the association of his mature life, and in feeling. He was the son of Richard Bache, for many years Postmaster of this city,—the grandson on the maternal side of Alexander James Dallas, and great-grandson of Doctor Franklin. Here, on the 19th of July, 1806, he was born, and here he was educated until he went to West Point, where he was graduated with the highest distinction, and was then commissioned as Lieutenant of Engineers. He did not remain long in the military service, his tastes being purely

> scientific. . . . But the great work of his life was the Coast Survey, of which he was Superintendent for many years. Under his guidance it weathered the political storms which, from time to time, threatened it, and was as nearly as may be completed. In former and happier days [before the war], the great opponent of the Survey, and incidentally of Mr. Bache, was Colonel Benton. Their advocate and zealous supporter was a statesman whom now it is heresy to praise. Alexander Dallas Bache in the Survey, and Montgomery Meigs in the great architectural and engineering work in the District of Columbia, owe everything to him who is now a prisoner at Fortess Monroe (*i.e.* Jefferson Davis). . . .

Here is a summary of the connection of Alexander Dallas Bache with the American Philosophical Society, although in describing this connection it has been necessary now and again to touch upon other organizations and topics assigned to other speakers. The life of Bache was so full of activity that it overflowed in many directions, but I think it may fairly be said that its well-spring was here in this Society which his great grandfather founded and which he honored so greatly through all his professional life.

ALEXANDER DALLAS BACHE AND HIS CONNECTION WITH THE FRANKLIN INSTITUTE OF THE STATE OF PENNSYLVANIA

HENRY BUTLER ALLEN

Secretary and Director, The Franklin Institute of the State of Pennsylvania

(*Read February 14, 1941, in Commemoration of A. D. Bache*)

Two years after the death of Alexander Dallas Bache, which occurred in Newport, 1867, a eulogy appeared in the *Journal of The Franklin Institute,* which ended with these words:

> It is rare that a man leaves us without a successor appearing to take his place, but the void which Mr. Bache leaves in the general scientific strength of the country has not yet been filled.

It was inevitable that a man of Professor Bache's caliber and scientific attainments should have become a leader in The Franklin Institute. He came to Philadelphia as Professor of Natural Philosophy and Chemistry in the University of Pennsylvania, when he was twenty-two years old. This was the year 1828. Four years before, the energy and enthusiasm of other young men had brought into being a virile young society—The Franklin Institute. The leaders who established the organization were Samuel V. Merrick, not quite twenty-one years of age, and Professor William H. Keating of the University of Pennsylvania, also a young man. Professor Keating had completed his scientific training in France and Switzerland and had returned to the United States full of zeal for the diffusion of scientific knowledge. He was immediately appointed to the newly created chair of Chemistry in its application to Agriculture and the Mechanic Arts, at the University of Pennsylvania. There can be no doubt that his enthusiasm for the new society was communicated to his colleague, Professor Bache. That a close association existed between the two men is evident from the minutes of the Institute, and when Professor Keating died in London in 1840, Professor Bache offered the resolution expressing the regret and loss of the Institute; and he served on the Committee appointed to condole with the family.

But we like to believe that family pride and the challenge of his ancestors played an important part in Bache's interests and activities in The Franklin Institute. Born in 1806, he was the son of Richard Bache, grandson of the illustrious Franklin for whom the young society had been named. Endowed with many of the attributes of his noted great grandfather, this young man would naturally contribute his talents to adorn and perpetuate so fitting a memorial to one of his own kin.

He was not the only one of that name to be connected with the Institute. A Richard Bache was elected a member in 1824, a Benjamin Bache is mentioned among the early members, and a Franklin Bache's name appears many times, particularly as a professor of chemistry in the schools of the Institute. It will be recalled that Sarah, Benjamin Franklin's daughter, married Richard Bache and became the mother of eight children. One, Benjamin Franklin Bache, evidently displayed traits of character which endeared him to his grandfather who took him to Paris, there to act as his secretary. The active and able young men just named from the early annals of the Institute were the sons of this favorite grandson of Franklin, and first cousins of the gentleman we are discussing today.

Alexander Dallas Bache was proposed for membership in The Franklin Institute by Samuel Vaughan Merrick, in April, 1829, and was duly elected the following month. His first public connection with the Institute was mentioned when it was recorded that on "March 25, 1830: A stated monthly meeting of the Institute was held in their Hall. Professor A. Dallas Bache was appointed Chairman."

From that time on, scarcely a month passed without some mention of the name of Alexander Dallas Bache. He served on the Board of Managers from 1831 to 1839, and was Corresponding Secretary from that time until 1843. His services to the Institute included membership in the Committees on Meteorology, Inventions, Instruction, Meetings, and Publication, and on the Committees on Patent Laws and the Manufacturing Establishments of Pennsylvania. In 1834, the Committee on Inventions was dissolved and in its place there was established the "Committee on Science and the Arts" continuing the work of the other committee but having enlarged powers and a wider field of labor. As originally constituted, membership in this committee was open to all members of the Institute in good standing who chose to enroll their names, and who by thus voluntarily associating themselves with the committee pledged themselves to perform the duties assigned to them. Bache was chosen first chairman of this Committee on Science and the Arts.

This volunteer system of service continued for fifty years, when an elective system was put into force which is the method used for membership in this Committee today. From the very beginning this Committee has performed a useful scientific service, as attested by its records of investigations of many subjects of importance. It is the branch of the Institute's work which perhaps more than any other illustrates the spirit which animated the founders and which their successors have worthily perpetuated and striven to improve and extend. It meets the urgent demand felt by inventors and discoverers in science—that some competent, trustworthy and wholly impartial body should exist to which they may turn for judgment and expert opinion on the usefulness of their researches.

The history of such a distinguished member of The Franklin Institute is necessarily the history of the Institute itself. Professor Bache was one of the Committee appointed by the Board of Managers to inquire into and report upon the probable causes of numerous explosions of boilers, and to find the best way to obviate the recurrence of such an evil. While research was still in progress, the Secretary of the Treasury of the United States requested that further extension of the Institute's inquiry include the prevention of steam boiler explosions. This led naturally to an investigation of the strength of materials, and the Committee devised apparatus of various forms for the testing of metals, building materials, steam boilers, etc. The results of this extensive work were published in the *Journal of The Franklin Institute* in various issues from 1831 to 1837.

Besides doing valiant work on the Institute's committees, Bache found time to contribute many important papers to its *Journal*, and to be interested in the exhibits of the Institute. It is recorded that he delivered the address at the close of the Annual Exhibition in 1842, an honorable appointment decided upon by a special committee.

Another important phase in the early days of The Franklin Institute, made possible by the character of its members, was the splendid co-operation which existed between the Institute and sister organizations. The meeting of a small group of men to decide whether or not The Franklin Institute should be brought into being was held in the "Philosophical Hall"; the Committee on Lectures in the early minutes reports that they "had the satisfaction of announcing to the Board that the Trustees of the University had granted the Society the use of their Academy on Fourth Street for the introductory lecture." (This was in April 1824, the Institute having been organized in February of the same year.) And in December, 1825, we find in our minutes the following:

> RESOLVED: That the thanks of the Board be offered to the American Philosophical Society for their liberality in leaving the busts and medals of Franklin in their possession for the use of the artist who prepared the medal die.

The Franklin Institute, having been aided and encouraged in its formative period, was quick to co-operate with others, and again we come back to Alexander Dallas Bache.

We find that upon his appointment to the presidency of Girard College he was commissioned by his Board of Trustees to visit Europe, in order to gain first-hand knowledge of the superior types of education which existed there. Dr. Bache gave two years to his travels. When he returned he presented a full report of his findings (*Education in Europe,* Report to the trustees of Girard College for Orphans, 666 pages, Phila., 1839). Appropriately enough, his report was reviewed in the *Journal of The Franklin Institute* (January, 1840). In 1850 we find the Institute sponsoring "The School of Design for Women," and in 1889, at the November meeting of the Institute, it is recorded:

> President Joseph M. Wilson, at the Stated Meeting of the Institute . . . made some remarks upon his observations of trade schools in France and England, having been charged with the task of investigating the organization and mode of operation of such institutions on behalf of the projected school of Mr. A. J. Drexel of Philadelphia.

So again, fifty years after Bache made his report for Girard College, the *Journal of The Franklin Institute* carried a report "On Schools: with particular reference to Trade Schools," which work resulted in the founding of Drexel Institute. The report had been made by Joseph M. Wilson, noted civil engineer and architect of his day, who designed buildings for the World's

Fair of 1876, and built many famous bridges. He was President of The Franklin Institute at the time he was commissioned by Mr. Drexel to gather data on trade schools.

The Institute co-operated with societies in other cities as well. Minutes of October, 1840, read:

> Professor Bache reported that he had visited the Fair of the American Institute in the City of New York in company with several members of the Institute, as a Delegate from the Institute, and gave an account of the Fair.

I have cited these instances just to show that the encouragement and co-operation received by The Franklin Institute in its early days was remembered and passed along when other institutions were thought of and needed publicity.

We have the pleasant conviction today that our building on the Benjamin Franklin Parkway, and the work carried on within its walls, constitute a fitting memorial to the great American whose name it bears, but an occasion such as this properly makes us pause and deliberate upon the great responsibility which we have inherited. We not alone must make the Institute worthy of its patron—we must live up to the high courage and lofty ideals of gentlemen such as Professor Bache, who, though called to serve his Government, never lost touch with the Institute. He freely gave of his time and his talents, so that, upon his death it is recorded:

> By the death of Professor Alexander Dallas Bache, the Franklin Institute has lost its most distinguished member, one to whose labors, while a resident of this city, it was indebted in an eminent degree, for the credit and reputation it obtained for its experiments on Water-power, on Steam Expansion, and on the Strength of Materials. . . .
>
> That the Institute will ever cherish . . . his generous and warm sympathy with its members in every movement that would add to the progress of Science and the Mechanic Arts and to the cultivation of liberal opinions and hearty social intercourse and good fellowship;
>
> That we mourn the loss . . . of so good and great a man!

Of such characters The Franklin Institute has been created.

THE CONNECTION OF ALEXANDER DALLAS BACHE WITH THE UNIVERSITY OF PENNSYLVANIA

EDWARD P. CHEYNEY

Professor Emeritus of European History, University of Pennsylvania

(Read February 14, 1941, in Commemoration of A. D. Bache)

Alexander Dallas Bache, then young and unknown, was appointed in the fall of 1828 to fill the position in the University formerly occupied by Robert M. Patterson, who had resigned to accept a professorship in the University of Virginia. His appointment came at a crisis in the University's life. Its affairs had been going badly; its income was inadequate, the number of students in the department of Arts was small, their attendance irregular, their behavior frequently bad. Much the same was true of the other colleges of the time. The early decades of the nineteenth century were a time of low water in academic life in America. The Trustees of the University of Pennsylvania not being able to obtain for the institution the success they wished, decided shortly after the resignation of Professor Patterson to reorganize the Faculty. This they did by vacating the chairs of all the remaining professors and appointing a committee to select a new Faculty. It was as one of the nominees of this committee that young Bache was chosen. The Trustees would gladly have reelected Professor Patterson but he preferred to accept the Virginia position. They then elected Henry Vethake, at that time an instructor in Mathematics and Chemistry at Dickinson College, but he declined though some years later he was elected Professor of Mathematics at the University and remained in its service for more than twenty years.

The committee then nominated Alexander Dallas Bache and he was elected Professor of Natural Philosophy and Chemistry, Sept. 16, 1828. Although but twenty-two years of age (men matured early a hundred years ago), the choice was a natural one. Bache was a Philadelphian. He was born here and educated at a local classical school (Claremont Academy), till appointed a cadet at West Point. He was, on his father's side, a great-grandson of Benjamin Franklin, a grandson of Richard Bache, husband of Franklin's daughter, who was active in the Revolution and Postmaster of the Colonies. On his mother's side he was a grandson of Alexander James Dallas, formerly Secretary of the Treasury of the United States, and a nephew of James M. Dallas, Vice President of the United States and later Minister to Great Britain. A cousin was Dr. Franklin Bache, a practicing physician in Philadelphia, long a prominent member and officer of the Philosophical Society. He and his relatives were a typical group of that intellectual-scientific society, as characteristic of Philadelphia in the early nineteenth century as was a somewhat more literary group of New England; but the history of whose flowering has not yet been written.

Precocious and studious, Alexander Dallas Bache graduated at West Point in 1825, at the age of nineteen, at the head of his class and with special commendation from the Secretary of War, who was himself present at his examination. He remained at the Academy for a year on an appointment to teach Engineering, then was detailed to assist in building some fortifications (Fort Adams) at Newport, Rhode Island. Here he remained two years. It is of some interest to note that during this period he made a scientific study and compiled a list of the shells of that neighborhood. When it is remembered that the historian Henry C. Lea began his long and productive career by studying and writing on marine and freshwater shells on which he had published in scientific journals ten papers before he published his first historical essay, one is led to wonder whether that study is not a good introduction to any other branch of learning. Certainly the accuracy of observation and description, the training in classification and the impartiality of treatment fundamental to all scientific study form an admirable preparation for any intellectual career. It was at Newport that in September 1828 he received from Nicholas Biddle, Chairman of the Committee of the Board of Trustees, the notification of his appointment in the University. His graceful letter of acceptance and appreciation signed as Second Lieutenant in the Engineer Corps United States Army still lies among the Biddle papers at the University. Like many another young teacher, on the strength of his appointment he married, resigned from the army, and October 1828 appeared at the University ready to take up his teaching. In the meantime the other vacancies in the little Faculty were filled and the next year a new College building was erected at Ninth and Market streets in the place of the old one at Fourth street; so it was in this reorganized and more hopeful University he took up his work.

His duties were arduous. He met each of his three classes every day of the week, including Saturday. In his second year he was made Secretary of the Faculty and a vast amount of writing in his small, clear hand remains to testify to this part of his College work. He kept the minutes of the weekly Faculty meetings, wrote letters transmitting reports to the Board of Trustees, and corresponded in the name of the Faculty with parents who protested against the disciplinary action taken against their sons. It is not a matter of surprise that he soon appealed to the Trustees for an assistant to do some of this routine writing. He proposed that some one should be paid $100 a year for this work, but after some discussion, at the suggestion of the Provost one of the students was relieved of his College fees in return for giving some help; but there does not seem to have been much diminution in the amount of his written work. As to his teaching, it gave an outline—it can hardly have been anything more—of the whole field of the physical sciences. His students were boys of sixteen, seventeen or eighteen years, corresponding in age to those of the last three years of our present high schools. But perhaps boys as well as their teachers matured more rapidly a hundred years ago. There were from twenty to thirty in each of his classes. The teaching of the physical sciences, "natural philosophy," as that group of subjects was called in distinction from moral philosophy—was an old tradition at Pennsylvania, which was less exclusively addicted to the classics than most other colleges; but occasion was taken of the general reorganization of 1829 and the accession of Prof. Bache to increase their share in the curriculum. According to the catalogues of this period he taught to the Sophomores "the Elements of Natural Philosophy and Chemistry," to the Juniors, under the head of Mechanics "the Doctrines of rest and motion as applied to solids and liquids and to machines"; under the head of Physics, "Electricity, Magnetism and Electro-Magnetism", and under the head of Chemistry, "Inorganic Chemistry." To the Seniors he taught "Optics as a branch of Physics," Astronomy, and the Steam Engine, and completed Chemistry. He lectured also on the Elements of

Geology and Mineralogy, and gave a general review of the whole field of Physical Science.

Beyond these brief titles in the catalogue and reports of examinations we have little knowledge of the content of these courses. None of his notes so far as known have been preserved. A few glimpses of his methods can be obtained. In the second year of his incumbency he petitioned the Trustees for a grant of $250 a year, $100 of it for materials for the teaching of Chemistry, $50 for current repairs to the physical apparatus, and $100 for technical assistance in preparing materials while he was lecturing. As he says, only with such provision can his subject be taught "with vigor and activity." The kind of teaching in which the students themselves should participate in experiment was still three-quarters of a century ahead. The Trustees seem to have been rather shocked at his request; it was referred to the committee on appropriations, but was granted then and yearly thereafter. In 1835 he finds he has overdrawn this account, but states that he had paid the excess himself. He apologizes for his carelessness but explains that he has always carefully deducted the cost of any material he has used in his own research, though he acknowledges that "in research I have felt much restricted by want of means."

Some idea of his methods of teaching may be obtained from the text-books he used. In Chemistry he used the well known work published in England by Edward Turner in 1827, the year before he began teaching. It was almost immediately reprinted in America and a second edition was published in Philadelphia in 1829. It was extremely popular and a sixth edition, edited by Dr. Franklin Bache, was issued in Philadelphia in 1840. One is somewhat surprised to find this text-book prescribed by Bache to be used with his elementary course in Physics as well as in Chemistry, until it is noticed that Turner prefixed to the chemical portion of his book an admirable chapter of some seventy-five pages on what he called "The Imponderables"—Heat, Light, Electricity and Galvanism. For his more advanced work in Physics he made use of the well known *Treatise on Optics*, by Brewster; but after using it in its English form for three years he himself prepared a new edition of this work which was published under his name as editor in Philadelphia in 1833. Although this contained various original notes, its principal characteristic was a long appendix devoted to the two subjects of the Reflection and Refraction of Light. It was distinctly an American college text-book. He says in his Preface that his object in undertaking the new edition was to adapt it for use in "those of our colleges in which considerable attention is given to the course in Natural Philosophy." America was coming of age in still another phase of her life. He claims that it is not an untried form of instruction; that most of it has "entered into the mathematical part of the course taught to the senior class in the University of Pennsylvania." He signs himself "Professor of Natural Philosophy and Chemistry in the University and one of the secretaries of the American Philosophical Society." In his last year at the University he used Gummere's *Astronomy.* For his other courses he seems to have relied largely on the popular works of Lardner in his *Library of Useful Knowledge.* In 1832 he translated from its French form and published in the *American Journal of Science* Berzelius' *Essay on Chemical Nomenclature*, but this was certainly too advanced for use with his college classes. In 1836, in recognition of the provision in the college building of a special room for the courses in Geology he presented to the University "as a nucleus for a geologic cabinet" his collection of some 400 or 500 specimens of minerals. In 1837 the Trustees acknowledged his attainments by conferring on him the degree of LL.D. He was still only thirty-one years old.

In the paucity of records concerning his methods of teaching it may be helpful to note that of all the cases of students' disorders that encumber the records of the time, none took place in Professor Bache's room. Professor Adrain, Professor Wylie and even the polished Professor Henry Reed frequently "arraigned," as the expression of the time was, students for inattention, disrespect or disorder in their class rooms, and even the Provost complained occasionally of bad behavior in chapel. But on the minute book there is not a complaint emanating from Professor Bache's room. Even his successor had trouble with his students, so it can hardly have been the fascinations of the subjects of physics, chemistry and geology that kept the students orderly. It would seem to be possible to make teaching so interesting that students will not want to misbehave.

Of the more personal side of his relations with his students there is more record than of the purely intellectual side. The traditional story of the origin of the Zelosophic Society describes a group of students in 1829 engaged in a dispute concerning the policy being pursued by President Jackson of removing his political opponents from office, when Professor Bache approached the group and was asked his opinion. This, with the usual professorial caution, he declined to give, but suggested that they hold a formal debate on the subject, intimating his willingness to act as moderator. From a series of such debates arose the new literary society in rivalry with the older Philomathean. The documentary evidence for this occurrence is somewhat slight but there is no doubt of the early and reciprocal interest between Prof. Bache and that Society. Long afterward the Society had a portrait of him painted, which now hangs in the office of the President of the University with the statement that it is the gift of the Zelosophic Society. There is other testimony to his popularity. At an Alumni dinner held in the year 1849, some years after he had left the University, William B. Reed offered a toast to Professor Bache and in speaking to it said: "I am a graduate of the class of 1832, the first class that commenced their studies and graduated under the new Faculty, of which Professor Bache was an efficient member. No one who has the good fortune to have been a pupil of Professor Bache can forget the amiable simplicity of his character, his devotion to the cause of science, his true and loyal attachment to the University and his warm and abiding friendship." What more could have been said of the friendly spirit of Mr. Chips? The same friendly relation with his students may be read between the lines of his recommendation to the Secretary of the Navy of one of his old students who was seeking in 1845 a position in the Navy, to whose talents he testifies and whom he describes as "an assistant at the magnetic observatory at Philadelphia."

When he resigned, the resolutions of regret adopted by his colleagues of the Faculty say that "their first and most pleasing recollection is suggested by the fact that the connection has been distinguished by an unbroken harmony." They speak of his "disinterested character and the amenity of their late association . . . which merits a distinct and peculiar record." Kind resolutions on parting must not be taken too seriously. Such testimonials are apt to be undiscriminating and even fulsome. But there have been members of Faculties of whom their colleagues could not possibly have spoken in just those terms. Even the resolutions of the Board of Trustees adopted long afterward, on news of his death, after a life of much distinction, call attention to the fact that while a Professor at the University he had not only "shed lustre on the institution by the variety and extent of his attainments" but had "secured the respect and regard of the students and of the Trustees."

The connection with the University which had begun in October, 1828 lasted for eight years, until July 21, 1836. On that day he was elected President of Girard College and on the same day resigned in a letter of regret his professorship in the University. The promptitude of this resignation was due to his desire that the Trustees should have plenty of time during the summer to choose his successor. It was accepted at a special meeting of the Board two days later. His successor, Lieutenant Roswell Park, was elected the next month and began his service at the University in September. The early career of Roswell Park had been curiously similar to that of Bache. He had entered West Point six years later, like Bache graduated at the head of his class, joined the engineer corps. He was at the time of his election to the University engaged on work on the Delaware Breakwater. His career at the University while following that of Bache in a general way, seems to have been colorless, and as time passed on he turned to the study of theology and in July 1842 decided to seek ordination and resigned his position. During this period of six years Bache had been engaged in educational work which will be described by another speaker. This work had now, however, ceased to occupy or satisfy him, and the Trustees of the University were glad to reelect him to his old position. This was in August 1842. I do not find anything especially characteristic of this second period of his connection with the University. It came to an end in little more than a year, though his name appears in the catalogues of both 1842–3 and 1843–4. His second period at the University had hardly begun however when in December, 1843 he was appointed head of the United States Coast Survey. He was embarrassed at leaving the University at a time so inconvenient to the Trustees, and wrote offering to hand over all his lecture notes to a successor or to serve without pay till a successor was found. The Trustees decided to give him leave of absence without pay for the remainder of the college year. He immediately accepted the government appointment and his connection with the University came to an end December, 1843.

A strict interpretation of my subject would bring it to a close at this point. But it would give but an incorrect impression of the life of Alexander Dallas Bache while he was connected with the University. His active life was twofold; he had scarcely begun his work as a teacher before he appeared as an habitué of these rooms and those of the Franklin Institute and the results of his research in various aspects of nature began to appear in their journals. He was elected a member of the Philosophical Society April 17, 1829, but six months after he had come to Philadelphia as a teacher in the University, and he appears as an appointee on a committee of the Franklin Institute three months later. Like his great-grandfather he was curious about all physical phenomena. Just as Benjamin Franklin utilized whatever time he could spare from his printing office, his duties as Postmaster, Clerk of the Assembly and his multitudinous public interests to experiment on electricity, the spread of oil on water, fireplaces, the direction of storms and a score of similar subjects, so Bache escaped from his teaching and secretarial work at the University to make observations of the variations of the compass, the periodicity of meteors, the amount of rainfall at different elevations and in a dozen other forms of physical phenomena. He published during the first eight years of his service at the University thirty-four papers reporting his observations on such subjects while the fifteen others published during the years from 1836 to 1843 must have been drawn largely from his researches during his teaching period and his studies abroad. I have no intention of describing or even enumerating these papers or estimating their addition to the sum of human knowledge. This is the task of those who are competent to do so and to whom the duty has been assigned. I would like, however, to draw from one of his earliest reports a few incidents that form a connecting link between his life as a teacher and as a scientific observer. This report was read before the Philosophical Society November 16, 1832, and is printed in the TRANSACTIONS of November, 1832. He says "During the months of August and part of September of this year the usual summer vacation of the University permitted my absence from the city, and finding myself favorably situated for meteorological observations I undertook to observe the diurnal fluctations of the barometer and ultimately the hourly variations of the horizontal needle." The "favorable situation" for observation he describes later as "about one mile from the village of Westchester, in Latitude 39°58′ and in Longitude about twenty-one miles northwest from Philadelphia." He describes with meticulous care the platform on which he placed his instruments "in a garden more than forty feet from the house and fifteen feet from a small paling. The garden was upon the side of a hill, the ground sloping toward a meadow; a hill enclosed this gorge both on the east and the west; behind the hill the sun passed about eighteen minutes before the time of setting." Here he made his observations from August 29th to September 8th. As the University then closed for the summer vacation on July 31st, it had apparently taken him a month to set up his apparatus and prepare for his daily and hourly observations. In these he had the help of his pupil and ultimate successor John F. Frazer and of his wife. These observations were all reported to the Society with a series of tables and graphs. However little these mean to the unscientific reader, I like to think of this young college professor hurrying away at the close of the term to the Delaware county hills, to his "garden on the side of a hill," to make rigorous observations on earth magnetism.

If every address ought to have a thesis as well as a subject, to illustrate some proposition as well as give some information and express some thought, I should like to assert here my conviction that a good observer is apt to be a good teacher, and *vice versa.* The opposite statement is often made. This is, I think, a delusion. Generally speaking the same qualities of intellectual curiosity, mental alertness and unflagging industry that make the good research man also make the good teacher, and *vice versa.* Of this I have long been convinced; Alexander Dallas Bache is a good example.

Once again, long afterwards, the story of Professor Bache's life touches the University. His death occurred in Newport, Rhode Island, February 17, 1867. February 19 a meeting of the scientific and learned societies of Philadelphia was held in the chapel of the University at which a committee was appointed to arrange for paying respect to his remains as they were taken through the city on their way to their burial in Washington. On the evening of the same day, a special meeting of the Philosophical Society was held at which Mr. Fraley reported the proceedings of the meeting at the University. The use of this hall was thereupon offered to the committee and in the room above this in this building his body was laid in order that citizens of his native city and others should have an opportunity to pay him honor. At the meeting of the Trustees of the University of March 5, 1867, the expenses incurred at his funeral were ordered paid and the resolutions of respect already quoted were ordered to be drawn up.

BACHE AS AN EDUCATOR

MERLE M. ODGERS

President, Girard College

(Read at Girard College, February 15, 1941, in Commemoration of A. D. Bache)

It is a great privilege and pleasure to extend to this distinguished company a cordial welcome on behalf of our faculty and the members of our Board.

Truly it is fitting that members of the American Philosophical Society and other friends of science should foregather at Girard College to celebrate the establishment here a century ago of the first magnetic observatory in North America. You may already have learned from your program what the observatory looked like.

Several decades ago a new Philadelphia public school was named for Alexander Dallas Bache. Sometime later, in 1898, a portrait of Bache was presented to the school with considerable ceremony. One of the speakers on this occasion was Professor George F. Barker, who filled the chair once held by Bache at the University of Pennsylvania. Referring to Bache's Magnetic Observatory at Girard College Professor Barker said: "Not only is there no trace of the building itself or any of its parts to be found within the walls of that institution, but there is even a considerable difference of opinion as to its exact location. No single spot in Philadelphia surpasses this in scientific interest."

Spurred by the investigation that had evidently preceded this statement, the Board of Directors of City Trusts met on the same day that the portrait was presented, considered "the importance of placing a suitable tablet to locate" the Magnetic Observatory, and referred the matter to a committee with power to act. The tablet could not be placed. Dr. Adam H. Fetterolf, who was at that time President of the College, reported to this Committee that no definite means of locating the site had been discovered.

A month later Dr. Fetterolf's report to the Committee furnishes a pleasant commentary on the value of human testimony: "We have not yet found the exact location of Professor Bache's Magnetic Observatory. We have had the opinion of men who were employed in it, of men who as College boys played around it, and of others equally familiar with the building, and they differ widely in their recollections of the locality." The Committee of the Board reported "progress," a word that is frequently a forecast that nothing will happen, and action was postponed for a time. Finally the matter was allowed to become a memory. Dr. McIlhatten, the Head of our Mathematics Department, recently interested the Bureau of Surveys of the City of Philadelphia in the problem of locating the site of the Observatory. Only two days ago I learned from Dr. McIlhatten that this Bureau and the United States Hydrographic Bureau were in agreement that its latitude and longitude, as given in Bache's preface to the three published volumes of Observations, would locate the Observatory at or near Stillman and Poplar Streets, just southwest of the present campus and outside it but not outside the southwest corner of the original Peel Hall tract on which the Girard Estate has built houses. Bache doubtless tried to locate his Observatory far from the proposed school site.

I spare the audience a detailed description of the wooden building with its copper nails, its brass hinges, its complete absence of iron, its foot-thick partitions and walls, and its double doors and windows, all affording protection for the work that was carried on inside. It may have had a plebeian appearance but its origin was patrician; for it was designed, according to the report of the stated meeting of September 20, 1839, of the American Philosophical Society, by none other than Thomas U. Walter, who was elected a member of the Society the following month. This young man, like his friend Bache, distinguished himself early in his career, designing at the age of twenty-nine the great Corinthian structure that has long symbolized Girard College. When its glory departed, the Observatory returned to the care of the carpenter who had built it. He moved it to another site, used it as a shop, and later razed it.

Bache was graduated from West Point at the head of his class on July 1, 1825, when he was not quite 19 years of age. For a year after he had become an army officer he remained at the Academy and had his first teaching experience as an assistant in engineering. Following this he was transferred at his own request to engineering service under General J. G. Totten at Newport, Rhode Island, where he met his bride. In September, 1828, he was elected Professor of Natural Philosophy and Chemistry at the University of Pennsylvania. When he assumed his professorship the following month, he was only twenty-two years of age; but he had entered and had left an honorable profession, he had taught a year, he had had two years of field experience, he had begun his career as a researcher, and on the strength of his new position he had married.

For almost eight years Bache remained at the University, contributing articles on scientific subjects to several journals and taking an active part in the affairs of the American Philosophical Society and of the more recently organized Franklin Institute. In 1833 he had been elected one of the Trustees of Girard College, for which provision had been made in the will of the French-born humanitarian, Stephen Girard. This association led on his thirtieth birthday, July 19, 1836, to his unanimous election by his fellow trustees as the first President of the College. At the same time the Board instructed him to make an unhurried trip to Europe to examine and report upon educational institutions abroad in order that he might have "useful information" at hand in his organization of the new school.

In his letter conveying to Bache the instructions of the Board Nicholas Biddle, its chairman, had set no limit to his stay in Europe. He says: "It is this anxiety that your investigation should be complete, which induces them not to fix at present any period for your return. . . . While, therefore, they are very anxious to open the College with the least possible delay, they deem it so much more important to begin well than to begin soon, that they postpone naming any limit to your stay in Europe, until you are able to apprize them of your progress." Little did either Biddle or Bache realize in the fall of 1836 that for a variety of reasons it would be over eleven years before Girard College would be opened.

It may well be that both as a scientist and as an educator Bache's work and Bache himself, because of the admiration felt for his great ancestor, were received in Europe with more attention than even his scholarship and personality merited. In Germany, for example, he visited an elderly scholar, who greeted him with an embrace, kissed him on either cheek, and exclaimed, "Mein Gott, now let me die, since I have lived to see with mine own eyes an emanation of the great Franklin!"

What a boon it would be today to those engaged in educational surveys if they were able to follow out the instructions of Bache's Board "to domesticate yourself, if practicable, in these institutions"! He left America in September 1836, studied more than 278 schools in the British Isles, France, Germany, Austria, Switzerland, Belgium, Holland, and Italy, and returned in October 1838. In the following winter he completed the writing of his monumental "Report on Education in Europe" and published it in the spring of 1839. No one could have made the study of European education under more favorable circumstances. Bache produced a report that was a masterpiece of educational research and compilation. A century later it is difficult to evaluate what effect, if any, it had upon American education at the time it was published. Shortly after Bache's death Professor Joseph Henry, looking back almost four decades, said of the book: "It has done more, perhaps, to improve the theory and art of education in this country than any other work ever published; and it has effected this not alone by the statement of facts derived from observation, but also by the inferences and suggestions with which it abounds."

Bache was well aware that Mr. Girard had "intended no ordinary Orphan Asylum to be created with the immense fund which his liberality intrusted to the authorities of the city" (*Report on Education in Europe*, p. 2) and that the founder had furnished the trustees, to quote from the conclusion of the Report, with "the means of establishing a series of model schools for moral, intellectual, and physical education, embracing the period of life from early youth almost to manhood, the importance of which to our city, and even to the country at large, can hardly be estimated" (*id.*, p. 606). But Bache was a man of energy and the long delay in opening Girard College made him restless and eager to put his talents to work.

In October, 1838, the month of Bache's return from Europe, the Central High School of Philadelphia, the first American public high school outside the New England states, had been opened with four professors and sixty-three students but no head. Here Bache saw his opportunity to be of service and to organize a school under the principles that he had so recently formulated. The buildings at Girard College were far from being ready, and though his salary was continued he had but little work to do. By obtaining the headship of the new school he could try out his ideas. In November of the next year, 1839, Bache was elected the first

President of the Central High School, at the same time assuming a kind of superintendency of all the public schools of the city.

Bache continued to use the facilities of the College for several years after he actually severed his official connection with it, for the magnetic observations were made here from 1840 to 1845, though it is to be noted that from Bache's removal from Philadelphia in December, 1843, they were under the direction of Professor John F. Frazer, who had been a student of Bache, a fellow worker in many of his scientific studies, a colleague on the faculty of the Central High School, and finally his successor on the faculty of the University when Bache entered the Coast Survey.

Bache spent three years in the public schools of Philadelphia, most of his time being devoted to the Central High School which he completely organized. As was to be expected, its curricula placed emphasis upon the sciences and training in the vernacular. Expensive scientific equipment was purchased, especially for the astronomical observatory. Several years ago Dr. Cheesman A. Herrick, who wrote the "History of Girard College" while serving as its president, said: "While Bache did not work out in detail the educational scheme for Girard College, we may accept the plan which he did work out for the Central High School as an indication of what he would have done, had he actually begun the work at the College. . . . Bache's influence on the Central High School has continued even to the present."

Dr. Robert Ellis Thompson, President of the Central High School from 1894 to 1920, says in an unpublished history of the school that Bache gave it character. "His influence upon the School," observes Thompson, "corresponded to his hearty, cheerful, courageous English temperament. . . . He illustrated the qualities he sought to foster. . . . He was loyal to West Point, and tried to convey to the School what good he had found there. He would have her graduates loyal to the Central High School, and advised them to form an association of her alumni, to keep her spirit alive in them by yearly meetings and addresses. In later years he constantly labored for her welfare, and in 1859 he delivered one of the annual addresses before the Alumni." Many years later the Associated Alumni of the High School commemorated Bache by endowing the Alexander Dallas Bache Physics Prize which consists of a substantial sum awarded to the member of each graduating class "whose average in physics shall be the highest."

On July 14, 1842, Bache presided at the first commencement of the school, and this was his last official appearance as its head. In the autumn of that year he returned to the University of Pennsylvania and to his former chair as Professor of Natural Philosophy and Chemistry, in which he remained for but little more than a year. In the latter part of 1843 he accepted the post of Superintendent of the United States Coast Survey and his formal career as an educator thus ended at the age of thirty-seven.

One of his admirers, Henry, says that in his Philadelphia days he was able to carry his regular duties, his research studies, and his committee work in the learned societies "by a division of his time into separate periods, to each of which was allotted its special occupation." Of Bache's University classroom work Henry says: "He was a zealous and successful teacher, to whom the imparting of knowledge was a source of unalloyed and inexhaustible pleasure. His pupils could not fail to be favorably impressed by his enthusiasm and influenced by his kindness. He always manifested an interest not only in their proficiency in study, but also in their general welfare. They regarded him with affection as well as respect, and while in other classrooms of the university disorder and insubordination occasionally annoyed the teachers, nothing was to be witnessed in his, but earnest attention and gentlemanly deportment. His success as an instructor affords a striking confutation of the fallacy which has not unfrequently been advocated in certain quarters, that men devoted to original research and imbued with habits of mind which it generates are not well qualified for the office of instructors."

As a good teacher Bache believed in the importance of good organization in a school. Referring in his Report (*Report on Education in Europe*, p. 84) to an institution in Hamburg, he says: "An efficient teacher may supply many deficiencies in a plan, but there are some cases, of which this is one, in which a defective organization places the remedy beyond the teacher's reach." It would appear that Bache, despite his experience in the immediate past as a university teacher, was capable of giving unusual attention to detail when he went on his educational mission though the foreign schools that he examined were on the elementary and secondary levels. Occasionally the details are so full as to be humorous, as when he mentions the fact that at one school the nine female servants have as one of their tasks "washing the feet of all the boys once a week" (*id.*, p. 17).

Bache knew that his Report could have a large usefulness. "If this account," he says (*id.*, p. 154), "should further contribute to awaken attention in our schools to improvements which have been introduced abroad, I am sure that the trustees of the Girard College will feel gratified at this useful result of their measures." Analogies to the Girard College to be established under the Girard Will must have suggested themselves: the avowed purpose of a school in Berlin that "not mere words should be taught to the pupils, but realities, explanations being made to them from nature, from models and plans, and of subjects calculated to be useful in after life" (as quoted in the *Report*, p. 517) must have recalled to Bache the words of Girard: "I would have them taught facts and things, rather than words or signs." Bache's eye from 1836 to 1838 was always upon the Girard College of the future (*e.g.*, *id.*, pp. 65, 149) with the thought that "we must raise our system upon the basis of the successful experiments of others, unless we would encounter the vexations incident to the acquisition of experience by our own failures" (*id.*, p. 170). "The trustees of the College," he comments (*id.*, p. 606), "have appealed to the experience of Europe to furnish data necessarily wanting in a new country, and it remains for them to apply the experimental deductions thus obtained from the old world with the vigour characteristic of the new."

By nature Bache was forward-looking and interested in educational experiments, new methods, and fuller attention to some neglected features of education, such as the playground that he calls the "uncovered school." Yet he recognized "the odium usually attaching to innovation" (*id.*, p. 365) and resented the fact "that what is an established system in one country, should by many be considered as an innovation of doubtful expediency" in another (*id.*, pp. 416, 505).

Like Quintilian and many other educators through the centuries, Bache made a strong point of individual differences and felt that differentiated curricula should be provided (*id.*, p. 270), that "different trains of instruction" should be pursued according to the students' "mental development and probable pursuits in after life" (*id.*, pp. 42–43), and that "different kinds of education are suited to different objects in life" (*id.*, p. 523). Along with others he anticipated questions of guidance (*id.*, pp. 379, 409, 658), the phase of education in which the individual approach is emphasized. He favored as a substitute for the formal report to parents a written statement of progress and conduct of the "anecdotal report" type that he found in a school in Berlin (*id.*, p. 248). He thought that the test by which schools are judged on the college success of their graduates was "essentially defective" (*id.*, p. 374) and also opposed the measurement of a school's worth by the conspicuous success of a few distinguished alumni (*id.*, p. 373).

Since Bache was interested in the individual approach he would, of course, oppose the system of mass education or mutual instruction associated with the names Lancastrian, Bell, and Madras (*id.*, pp. 51, 94, 171, 175, 182, 194, 196, 206) and he considered this type of instruction "a very unadvanced grade of public education" that was inexpensive but desirable only where nothing better was possible. He thought monitorial systems were doomed to have only limited success. He also opposed advanced forms of student self-government (*id.*, pp. 394, 413), though he felt that boys must be made to assume responsibility within reasonable limits (*id.*, p. 78) and endorsed the performance of chores by students (*id.*, pp. 31, 55) referred to in our day as the self-help plan.

The inescapable problem of discipline challenged his attention, to be sure. Bache thought that it ought to be based on "the Christian law of love . . . in school, if we would have it practised in society." He disapproved of the fagging system and the demerit system, used "mild" as a word of commendation in relation to discipline, and disliked useless regimentation (*id.*, pp. 28, 76, 118, 131, 139, 187, 262–3, 394, 507). In his Report to the Committee of Controllers of the Public Schools on the Organization of a High School for Girls, he sums up (pp. 9, 10) his attitude toward discipline: "The authority of the teacher being a portion of that of the parent, and delegated for the time being to him, the parental relation should, as far as possible, exist in a school."

Bache's ideas of moral education were also in line with those expressed in the Will of Stephen Girard (*e.g.*, *Report on Education in Europe*, pp. 128, 165, 188, 402, 403, 514). In moral education and in the more formal Biblical instruction Bache searched for whatever might be best adapted in Girard College. In Holland he found "an important experiment in communicating religious without sectarian instruction" (*id.*, p. 171; also pp. 88, 160, 180, 214–5), which he would doubtless have employed had circumstances permitted.

In discussing a school in Berlin Bache says: "The methods of instruction in this school are, in general, most excellent, and I was particularly struck with the small number of text-books employed. This is not peculiar, however, to this establishment, but is a feature in every good school in Germany. The master is expected to be so fully imbued with his subject, and expert in his art, as to be able to impart knowledge principally orally to his pupils, and in such a way as to adapt it to each individual" (*id.*, p. 241). Bache admired the German schools of a century ago. He liked their well-developed interest in scientific study, their program and methods of instruction, and their disciplinary systems. The *gymnasia* were apparently not strongly regimented, and the supposed defects of their teachers (*id.*, pp. 460–1) are interestingly enough the defects that are charged today against American college and university teachers.

Though the training of young children was foreign to his own earlier career, Bache found great interest in "infant schools," the best of which were at Glasgow, Edinburgh, and London (*id.*, p. 159), and in elementary schools, which in Great Britain strangely enough were inferior to those in Germany and elsewhere (*id.*, p. 170). At all events, Bache realized the large importance of elementary education (*id.*, p. 64; also pp. 154, 157, 164, 197, 293, 302). He was aware of the forces at work in this field, especially Pestalozzi (*id.*, pp. 94, 211, 213–4, 234, 287, 325, 360–1, 403), and he sought up-to-date elementary methods (*id.*, p. 105).

So far as Latin and Greek and higher mathematics as the core of all secondary education were concerned, Bache was skeptical, saying: "I am far from being one of those who undervalue classical culture, but I am convinced that to be at all effective it must be thorough, that it cannot be thorough when the instruction is terminated at an early age, and that there are certain minds very little or not at all improvable by language, as there are others similarly related to mathematical studies" (*id.*, p. 22). He thought "the amount of intellectual culture to be gained by the exercises of arithmetic and algebra . . . to be undervalued" (*id.*, p. 387) but the aptitude and the probable future career of the student had to be studied before he should consider entering upon higher mathematics. In believing that a modernized classical course was needed, and that "the exclusion of all, or nearly all, that characterizes modern civilization, brings discredit upon the system, and the worst foes of the legitimate use of classical culture are those who profess to be its best friends" (*id.*, p. 400),—in believing this, he saw eye to eye with Thomas Arnold.

The Report, written just a hundred and two years ago and picturing education at that time, reveals what great changes in curricula and subject syllabi a fast-moving century has brought about. This can be traced in the history of the Central High School, though I shall mention only the matter of choice of courses. Bache had given Central a choice of three courses, and even this arrangement probably represented a compromise with his plans. "It was not so much what he wanted, as what he saw he could get," says Dr. Thompson. The latter states that in succeeding presidencies the school sank "into a uniformity of curriculum which compelled every student to take the same course." This uniformity ended with Dr. Thompson's predecessor, Henry Clark Johnson, who was President from 1888 to 1893. Johnson's plan gave five choices. This liberality of selection was doubtless the main reason for Johnson's administration's being "regarded as a period of sweeping innovations which threatened the historical identity of the school. As a matter of fact, the changes made in his term were chiefly of restoration of what had been established by Professor A. Dallas Bache in the beginning." Truly Bache "gave character to the High School."

In passing, it is interesting to note that at Central Bache had built up a school library of over one thousand volumes, certainly one of the earliest and one of the largest school libraries of a century ago. Bache had previously built up for Girard College a library of the same size. It is a pity, indeed, that circumstances did not permit him to leave his influence upon Girard College. If the civic authorities and more members of its early boards of control had had the cosmopolitan spirit and the broad vision of Bache, a spirit and a vision that Girard himself would have applauded, Girard College would not have waited a half century to take its rightful place among America's great schools and it would long have enjoyed a wide reputation commensurate with its size and resources and the importance of its work.

When Bache's name was urged upon the Secretary of the Treasury for appointment to the vacant Superintendency of the Coast Survey, the Secretary thought that he was without administrative ability and referred to him as a "mere college professor."

In a few months following Bache's appointment, the Secretary had completely changed his tune and had become his firm friend and supporter. Bache was young in years but not in experience or in the depth and breadth of his thinking. The same qualities that made him a successful university teacher, a profound student of current education, an organizer and administrator of recognized ability in the public schools, a leader in scientific societies, and a distinguished researcher in the physical sciences led him to develop his associates as an understanding educator would and with them to make the Coast Survey a department of marked distinction.

One of his associates in the American Association for the Advancement of Science, Benjamin Apthorp Gould, speaks of Bache's "keen appreciation of humor, his love of pleasantry and jest, and his social geniality. . . . He knew the secret of obtaining work from his subordinates, by doing more than they did." He worked with each investigator, "stimulating his zeal, encouraging his hope, suggesting new ideas or infusing needful caution." Could anything better be said of a man who was both scientist and humanitarian than was said of Bache by Gould: "He had studied men . . . as he would study physical phenomena."

ALEXANDER DALLAS BACHE AS SUPERINTENDENT OF UNITED STATES COAST SURVEY 1843-1867

REAR ADMIRAL LEO OTIS COLBERT
Director, United States Coast and Geodetic Survey

(*Read February 14, 1941, in Commemoration of A. D. Bache*)

If "Professor" Bache—as he appears to have been generally called by his contemporaries—could return today he would doubtless be amazed by some of the latest methods of the Coast and Geodetic Survey, such as echo sounding and radio acoustic ranging, but he could rightly feel that these and many other similar developments were only the logical outcome of the work started by him almost a century ago. His readiness to investigate and adopt new methods was perhaps his outstanding characteristic, and one of the principal reasons for his success in the early stages of applied science in this country.

Originally recommended to Congress by Thomas Jefferson in 1807, the Coast Survey was not put into effective operation until 1832, under its first superintendent, Ferdinand R. Hassler. Numerous causes led to the adoption of an act completely reorganizing the Survey in March 1843, and upon the death of Hassler which occurred that same year it was Alexander Dallas Bache who was called upon to carry out the purposes of this Act, and to place the Survey upon a sound and steadily expanding basis.

It was natural to look to Philadelphia and particularly to select a descendent of Benjamin Franklin for this task. According to Joseph Henry's eulogy of Bache, his appointment was first suggested by the American Philosophical Society, and it immediately gained the support of the army (Bache was a graduate of West Point), of many educators, and of scientific institutions. Bache himself is said to have taken no part in securing the appointment, for while it opened a wider field for the exercise of talent and the acquisition of an enviable reputation, it also involved difficulties and responsibilities of the gravest character. In the words of Joseph Henry, "it would appear as if the training and acquisition of every period of his life had been especially ordained to fit him in every respect to the difficulties of this position. Besides the qualities we have enumerated, he possessed rare executive ability, which enabled him to govern and guide the diverse elements of the vast undertaking with consummate tact and skill. Quick to perceive and acknowledge merit in others, he rapidly gathered around him a corps of men eminently well qualified for the execution of the tasks which he severally assigned them."

Since, then, only a beginning had been made under his predecessor—although a beginning which Bache himself repeatedly acknowledged as having laid the foundation of the work—it was Bache who built up the Survey on that foundation, and led it steadily through the formative years when there was at times much opposition to it, and many attacks were made in Congress upon its usefulness and desirability.

The survey of the coasts accomplished up to 1844 extended only from New York Harbor to Point Judith in Rhode Island, and southward to Cape Henlopen. Bache's first major move was to divide up the entire coast line into sections, in each of which the essential operations were to be performed simultaneously by separate parties. Apart from the merit of this plan for the accomplishment of the work, it also possessed the great advantage of gaining support from all sections of the country along the coast, and at times when the Survey was under attack in Congress it was often the Senators and Representatives from the southern states who were its strongest supporters and advocates—a notable example being Jefferson Davis, while Senator from Mississippi in 1849. From operations in 9 states in 1844, the work was expanded to 16 states in 1845, while in 1848 parties worked in every state on the Atlantic and Gulf of Mexico, and that same year the first party set out with instruments on the long journey to the Pacific coast. Bache's administration thus saw the extension of the coastal surveys from the original small nucleus to newly acquired Texas and the western coast, while the purchase of Alaska took place in the year of his death. He is reported to have been in the habit of answering the question, often propounded to him by members of Congress, "When will this survey be completed?" by asking them in turn, "When will you cease annexing new territory?"

The great financial upheaval of 1857 led to economy in government expenditures, and in his report for 1858 Bache gave some interesting figures to justify the continuation of the Survey. From 1843 to the middle of 1857, the total cost was less than $4,250,000; appropriations had increased two-fold, but work executed had increased three-fold; the personnel in the Coast Survey was only 6/10ths the number in the English Survey. The president of the Royal Geographical Society of London stated in 1850: "I have studied the question closely, and do not hesitate to pronounce the conviction, that though the Americans were last in the field, they have leaped into the very first rank." A year later the same president speaks of the work of the Survey as "one of the most perfect exemplifications of applied science in modern times."

The annual report for 1859 contains the report of the "Committee of Twenty" of the A. A. A. S., appointed in 1857 to examine into the character and progress of the Coast Survey. The membership of this committee is a roster of the most distinguished educators and scientists of the period, and its twelve conclusions were in the main highly laudatory of the work accomplished by the Survey, as will be seen from the following excerpts:

> Their scientific value is witnessed, in the instance of the American Survey, by the spontaneous tributes of approval frequently and freely bestowed upon it—no less in regard to the ability, energy and skill displayed in its management than to the magnitude, variety and oftentimes curious interest of the results it has wrought out—by individuals and organized bodies of men whose high position as scientific authorities renders their opinions upon subjects of this nature entirely conclusive.
>
> This work has conferred many valuable benefits upon science, indirectly and incidentally, in the invention or perfection of instruments, in the improvement of methods of observation or computation, in the development which it has given to special subjects of interesting inquiry, and in the stimulus which it has furnished to the scientific talent of the country, especially in the field of astronomical observation and investigation.
>
> A careful study of the progress made from year to year, especially since the enlargement of the scale of operations under the present Superintendent, affords ample evidence that the work has been expeditiously prosecuted, and that the amount accomplished up to the present date is materially greater than has ever been accomplished in any other country in the same length of time and with the same means.

So much for the broad lines of the growth and expansion of the Coast Survey under Professor Bache from 1843 to 1865. During the last two years of his incumbency, while he was still nominally Superintendent, until his death on February 17, 1867, the work was in charge of J. E. Hilgard, who later became Superintendent in 1881. During these two years Professor Bache endeavored to regain his health, and spent some 18 months in Europe but without success.

Bache had the same type of inquiring mind possessed by his very distinguished great-grandfather, but at the same time was critical and sound in judgment, and is credited by Joseph Henry with having saved the Government several million dollars by his work on a Civil War committee charged with the examination of newly proposed inventions and projects to improve the art of war, "by preventing the adoption of plausible though impracticable propositions from which nothing but failure and loss could have resulted."

Some of the special fields in which Bache's administration proved eminently successful were the following:

(*a*) Determination of latitude and longitude; in 1851 he presented a list of 3240 geographical positions, with the comment: "It will be found that improved methods of determining the latitude by observations have been employed; that the various methods which astronomy and geodesy furnish for determining longitude have been applied; and that, in addition to the usual ones, we have introduced for the *first time*, as part of the geodetic work, the determination of the difference of longitude by the *telegraph*." The first use of the telegraph was in 1846, at the time of the opening of the first commercial line between Washington and Baltimore. The director of the Cambridge Observatory said in 1850: "This method of adapting the electric current to the wants of the astronomer, and which has grown up under the fostering care of the Coast Survey, is purely *American*, and is acknowledged as such in Europe."

In his eulogy of Bache delivered in 1868, Professor Benjamin Apthorp Gould declared that the Talcott method for latitude observations was then known abroad as the Coast Survey method; that in Germany in 1868 astronomers had just reached the excellence in longitude determinations equal to that attained by the Coast Survey ten or fifteen years before; and that the latitudes and longitudes of the Coast Survey primary stations were determined with more accuracy than the same co-ordinates were known for any European observatory.

Bache in his report for 1858 stated that the "American" method of longitudes "will bear favorable comparison, it is believed, with any other method yet devised. It requires only the electric connection with Europe to make its application to the first meridian of the survey complete." The actual use of the transatlantic cable for this purpose did not occur until just after Bache's death, but he had made all the preliminary arrangements and had sponsored the idea from the beginning.

(*b*) It is of interest to note that from 1844 to 1860, with two exceptions, all of Bache's annual reports to the Secretary of the Treasury were dated from places away from Washington, ranging from Maine to North Carolina. Prior to the Civil War he spent a considerable part of each year in the field for reasons of health, participating in the actual work of a field triangulation party and directing the main operations of the Survey by correspondence. His inability to continue this practice of storing up health, because of the pressure of war work, is believed to have hastened his death.

(*c*) Bache's administration saw the general introduction of steam navigation and he very early—in 1844—urged upon Congress the desirability of employing steam vessels for the work of the Survey. In 1847 he reported that a steamer's performance in sounding operations exceeded by fifteen times the best previous work accomplished by a sailing vessel. As early as 1850 a new steamer for use in the West Coast survey activities was under construction. With the further adoption of iron vessels, one of Bache's incidental tasks was membership on a committee charged with devising means of compensating compasses on such vessels.

(*d*) Many studies in oceanography were made under his direction. An extensive investigation into the Gulf Stream was started in 1844, and its structure and laws were detected for the first time. Agassiz was twice sent to Florida to study the growth of coral reefs. By 1860 fourteen sections of the Gulf Stream had been surveyed, 300 positions occupied and 3600 observations of temperature made. Tides and currents also received much attention, leading to the adoption of new methods and instruments, and the beginning of the publication of tide tables in 1855. A new way to carry on deep-sea soundings and to bring up samples of the ocean bottom was also developed. Special investigations under the direction of the Survey were conducted in the harbors of New York and Boston in cooperation with the local authorities.

(*e*) One of Bache's earliest aims was to popularize and render available to the public the utilitarian side of the discoveries of the Survey. In 1844 he first mentions arrangements for making maps and charts available for general use, at a cost sufficient only to cover the cost of paper and printing, "lower than that of the British Government and quite as low as that of France." That all possible aids to navigation were needed is shown by the figures for one year, 1855, when about 430 persons lost their lives along our coasts, and property damage amounted to $11,850,000.

When it was found difficult to secure enough skilled engravers he sent a representative to Switzerland in 1855 to engage a number of experts; the report for 1851 has a lengthy article on electrotyping operations of the survey, with apparatus and processes then entirely new; a year later the methods used for printing from lithographic transfers were described at length; and by 1866 an attempt was made to print maps and charts in colors.

Before the middle of the last century the sailing directions or coast pilots for the Atlantic Coast had been developed and published by private individuals or firms. As early as 1849 Bache started the compiling of such books to be put out by the Government as official documents, and the approaching completion of the sailing directions for the eastern coast was mentioned in 1857. Professor George Davidson's first edition of the Pacific Coast Directory was published in that year. The present comprehensive set of volumes of the Coast Pilot for all the coasts of the United States grew out of these early efforts instituted by Bache.

(*f*) From the outset Bache devoted considerable attention to his joint duties as Superintendent of the Office of Weights and Measures—the early ancestor of the present National Bureau of Standards. In the earliest days of the Survey this office was considered as of prime importance, and in 1832 its activities were much greater than those of the Coast Survey. The early functions of this office comprised mainly the preparation of standard weights and measures for distribution among the various states, together with methods for collecting duties on distilled spirits and similar activities.

(*g*) Bache's very great contribution to the study of magnetism will be described by another speaker on this program. In the field of geophysics, it is of interest to note that in several of his annual reports (1854 and 1862) there is mention of tidal waves on the west coast, believed to have been due to distant earthquakes or submarine volcanic explosions. In 1860 the Coast Survey sponsored a special expedition to the coast of Labrador to observe a total eclipse of the sun. In 1856 mention is made of the rapidly progressing library of the Coast Survey. History and geography were not overlooked, as Bache fostered several studies of the history of the early discoveries and explorations along both the Atlantic and Pacific Coasts, and in 1856 he wrote at length upon the difficulty of deciding upon the correct form of geographic place names, especially on the West Coast. In other words, Bache's outlook was so broad and his interest so catholic that any aspect of the work of the Survey received his enthusiastic encouragement and support.

(*h*) It would be inappropriate to omit from this brief statement a tribute to Bache's success as an administrator. In 1850 he emphasized the value of the experience and knowledge of the Coast Survey staff and officers, and their most cheerful, thorough, and zealous cooperation in every part of the work. He repeatedly endeavored to secure more adequate compensation for them, and with the exception of one incident near the beginning of his term of office—in which the official concerned was a disappointed rival of Bache for the appointment as Superintendent—he seems to have had no friction in his organization nor to have had any difficulty in securing and holding the willing and loyal devotion of his staff. One of his brothers, Lieutenant George M. Bache, U.S.N., serving in the Survey on a temporary assignment, was lost at sea in 1846, being washed overboard during a hurricane from the Survey brig "Washington," while engaged in work off the coast of North Carolina.

It was only natural that a man of Bache's character and achievements should be called upon for many special tasks related to but not a part of his principal official duties. Among the most important was his nomination in its act of incorporation as one of the regents of the Smithsonian Institution, in which office he was continued for twenty years until his death. In the words of Joseph Henry, long the Secretary of the Smithsonian, "to say that he assisted in shaping the policy of the establishment would not be enough. It was almost exclusively through his predominating influence that the policy which has given the institution its present celebrity was, after much opposition, adopted." As vice-president of the United States Sanitary Commission (the forerunner of the American Red Cross) he was very active in suggesting means for securing contributions and for their wise expenditure with a view to ameliorating the condition of soldiers and prisoners during the war.

He was also a member of the first commission appointed to examine into the methods of improving and expanding the light house service of the country, and upon its establishment became a member of the first Light House Board. Particularly during the Civil War he was called upon to serve on many special committees dealing with subjects in which he was especially competent to give counsel and advice—as, for example, in the adoption of the plans for League Island Navy Yard. One of the last major acts of his official life was a manifestation of his devotion to his native city, where he had hoped to resume his residence upon his retirement to private life. At the request of the Governor of Pennsylvania, he planned lines of defense for the city of Philadelphia, and partly superintended in person their construction. Joseph Henry attributes to this arduous task and to the exposure it involved the final breakdown of his health, weakened by years of overwork, from which he was not to recover.

Many tributes to and eulogies of Bache were written and published after his death. To quote Joseph Henry again, his valuable life may be said to have been one of the sacrifices offered for the preservation of the Union. Perhaps the most fitting eulogy to repeat here is that of his successor as Superintendent of the Coast Survey, Professor Benjamin Peirce of Harvard University, the most eminent American mathematician of his day, who in his first annual report for 1867 wrote the following:

> On the 26th of February, 1867, I received the commission of Superintendent of the Coast Survey. This important service originated with Hassler; but it received its efficient organization from Bache. . . . It is only necessary conscientiously and faithfully to follow in his footsteps, imitate his example, and develop his plans in the administration of the Survey. To describe what the Superintendent should do is simply to describe what Bache actually performed. . . . The ease with which the survey has kept pace with the progress of art and science, contributing more than its full share, so that its many processes are regarded as the unsurpassed models of their kind, has most arisen from the sympathy of my predecessor for all that was new, and his acute discrimination, which retained the improvement and rejected the fallacy. . . . I have before me the inspiration and example of my friend Bache. It is his organization. I have only to administer as he showed the way.

ALEXANDER DALLAS BACHE, A FOUNDER, FIRST PRESIDENT AND BENEFACTOR OF THE NATIONAL ACADEMY OF SCIENCES

FRANK B. JEWETT

President, National Academy of Sciences

(*Read February 14, 1941, in Commemoration of A. D. Bache*)

THE name of Alexander Dallas Bache does not raise very vivid memories in the minds of physicists of the present day. Quite unlike that of Joseph Henry, a close associate, his researches were concerned with phenomena which have not fructified into topics of either great practical or theoretical interest. His contributions dealt largely with terrestrial magnetism, magnetic surveys and problems in weights and measures—all of recognized fundamental importance, but nevertheless in departments of physics which neither then nor later could be regarded as spectacular or especially productive of expanding knowledge.

Yet not alone American science, but the American nation as well, is indebted to Bache. He was a pioneer on this side of the water in the field of scientific organization. Doubtless his interest in this sphere was greatly whetted by an opportunity accorded him by the trustees of the newly founded Girard College to visit Europe and study educational systems for the information of the board. We may safely surmise that it was from this extended visit and survey, in which he had occasion for discussion with many of the leading scientists of the old world, that Bache came to a clear appreciation of the value of an agency to advise the national government on scientific questions. During the course of his European tour he undoubtedly became familiar with the official governmental functions of such groups as the Royal Society in England and the Academy in France, so that then were sown the seeds later to ripen into our own National Academy of Sciences.

We can imagine that the need of such an institution in the United States formed a frequent topic of discussion between Bache and such close associates as Joseph Henry, Louis Agassiz, Asa

Gray and Benjamin Silliman. Thus, twelve years before the founding of the Academy, Bache delivered in Albany in 1851 an address as retiring president of the American Association for the Advancement of Science in which he dwelt at length on the need of a national scientific organization of the kind the Academy was destined to be. The following excerpts from that paper seem worthy of quoting, even today.

> . . . But first a few observations on the ordinary modes of promoting science; in connexion with which, I would throw out for your consideration some reasons which induce me to believe that *an institution of science, supplementary to existing ones, is much needed in our country, to guide public action in reference to scientific matters.*
>
> It is, I believe, a common mistake, to associate the idea of academical institutions with monarchial institutions. We show in this, as in many other things, the prejudice of our descent. We have among us the two extremes of exaggerated nationality and of excessive imitation: let us modify each by the other, and be wise. A national institute is not necessary to Great Britain, with her rich and powerful universities. Republican France has cherished her Institute, seeking rather to extend than to curtail its proportions. One of the most ardent of republicans is its perpetual secretary. . . . Nor does the idea of a necessary connexion between centralization and an institution strike me as a valid one. Suppose an institute of which the members belong in turn to each of our widely scattered States, working at their places of residence, and reporting their results; meeting only at particular times, and for special purposes; engaged in researches self-directed, or desired by the body, called for by Congress or by the Executive, who furnish the means for the inquiries. The detail of such an organization could be marked out so as to secure efficiency without centralization, and constant labor with its appropriate results. The public treasury would be saved many times the support of such a council, by the sound advice which it would give in regard to the various projects which are constantly forced upon their notice, and in regard to which they are now compelled to decide without the knowledge which alone can ensure a wise conclusion. The men of science who are at the seat of government either constantly or temporarily, are too much occupied in the special work which belongs to their official occupations, to answer such a purpose; besides, the additional responsibility which, if they were called together, they must necessarily bear, would prove too great a burthen, considering the fervid zeal, and I might almost say fierceness, with which questions of interest are pursued, and the very extraordinary means resorted to to bring about a successful conclusion. . . .
>
> Our country is making such rapid progress in material improvement, that it is impossible for either the legislative or executive departments of our Government to avoid incidentally, if not directly, being involved in the decision of such questions. Without specification, it is easy to

see that there are few applications of science which do not bear on the interests of commerce and navigation, naval or military concerns, the customs, the light-houses, the public lands, post-offices and post-roads, either directly or remotely. If all examination is refused, the good is confounded with the bad, and the Government may lose a most important advantage. If a decision is left to influence, or to imperfect knowledge, the worst consequences follow.

Such a body would supply a place not occupied by existing institutions, and which our own is, from its temporary and voluntary character, not able to supply.

Nothing, however, appears to have been done in political circles about Bache's recommendation until 1863, and it may well have been the crisis of the Civil War which furnished the needed emphasis to bring official action. In addition to the scientific group which discussed the national needs in science, frequently gathering at Bache's house on Twentieth Street in Washington, another visitor was Senator Henry Wilson. Perhaps it was Wilson who supplied a ready political liaison.

Out of the deliberations of this small group arose a decision to create at first a scientific commission, and the then Secretary of the Navy, early in February 1863, appointed three members: Bache (Superintendent of the Coast Survey), Joseph Henry (Secretary of the Smithsonian Institution) and Charles H. Davis (Chief of the Bureau of Navigation, Navy Department), to report on various "matters of science and art." These experts considered many subjects and gave their opinion regarding them.

Curiously enough, this so-called Permanent Commission was destined to pass virtually out of existence a few weeks later, for it was toward the end of the same month of February (1863) in which the Permanent Commission was appointed that a bill to incorporate the National Academy of Sciences was brought before Congress. The bill was passed March 3, 1863, and was approved by President Lincoln on the same day. As the leading scientific men in Washington at the time, Bache and Henry no doubt framed the provisions of the bill with the aid of such others as Charles H. Davis, Louis Agassiz, and Senator Henry Wilson, and in this act framed the charter of the Academy. Bache's name appears fourth in the alphabetical listing of names and states of residence in the Act of Incorporation, and is followed by the designation, "at large"—one of four or five so designated.

If time permitted I should like to discuss at length the masterly skill that the wording of the Charter exemplifies. Aside from a recital of the names of the original fifty members, it is a document of less than two hundred words. One section seven lines long provides for reports to Congress but gives the Academy complete independence concerning operating rules and the election of members.

Another section, of less than a hundred words, provides that the Academy shall hold an annual meeting at some place of its own choosing within the United States. It also imposes the obligation that ". . . the Academy shall, whenever called upon by any department of the Government, investigate, examine, experiment and report upon any subject of science or art . . ." the actual expenses to be paid by Government but the Academy to receive ". . . no compensation whatever for any services to the Government of the United States."

It is difficult for me to conceive of a more effective structure for insuring that at all times Government be in position to avail itself of the most competent unbiased assistance which the departments of knowledge represented in the Academy can provide. So long as the eminence of its membership is maintained, the Congressional Charter provides an almost impregnable citadel for disinterested service to the nation. It is a shield against the pressures of expediency, influence and the thousand and one things which so frequently bedevil and thwart our group undertakings.

During the temporary organization in April 1863 of the incorporators as members of the Academy, it was the feeling, expressed in action by the Academy, that each member should take an oath of allegiance to the United States of America and to the Academy. The form of oath agreed upon and administered to each member is quoted below:

> I, . . . , do solemnly affirm that I have never voluntarily borne arms against the United States since I have been a citizen thereof; that I have voluntarily given no aid, countenance, counsel, or encouragement to persons engaged in armed hostility thereto; that I have neither sought nor expected to exercise the functions of any office whatever, under any authority or pretended authority in hostility to the United States; that I have not yielded a voluntary support to any pretended Government, Authority, power or constitution within the United States, hostile or inimical thereto. And I do further affirm that to the best of my knowledge and ability I will support and defend the Constitution of the United States against all enemies, foreign and domestic: that I

will bear true faith and allegiance to the same; that I take this obligation freely, without any mental reservation, or purpose of evasion; and that I will well and faithfully discharge the duties of a member of the National Academy of Sciences. So help me God.

After Senator Wilson had opened the first meeting, in April 1863, Joseph Henry was asked to serve as chairman *pro tem.* of the Academy during the organization proceedings. At the same meeting, following the preliminary steps of organization, the Academy proceeded with the nomination and election of officers. Alexander Dallas Bache was the only nominee for the office of president, and on the first ballot was unanimously elected first president of the Academy. He was particularly qualified to hold this office, for he was cognizant of all the considerations which had led to the Academy's inception.

As the first president and as a member of the organizing committee, Bache had to do also with the framing of the constitution and by-laws under which the Academy was to operate as the scientific advisory body to the Federal Government.

The Act of Incorporation provided for the division of the membership into classes. During the first meeting of the Academy in New York, a committee on organization was appointed, and Bache was one of those selected for membership on the committee. In accordance with a mandate from the Academy, the committee divided the membership into classes. Class A, Mathematics and Physics, was further divided into Section I (Mathematics) and Section II (Physics). To this latter Section of Class A, Bache was assigned, and he remained a member of this Section until his death in 1867.

One of the first (if not the first) government committees appointed in the Academy was the Committee on Weights, Measures, and Coinage, to advise the Treasury. Bache was undoubtedly the best qualified member of the Academy to serve on this committee, but as President of the Academy he would not appoint himself. The Academy took matters into its own hands, however, and voted unanimously in favor of his accepting membership on the committee, and as he could not do otherwise, he accepted, and remained a member of the committee until his death.

Although long in failing health, Bache attended and presided at the meetings faithfully. His interest never waned during the few years which elapsed before his death in 1867.

In the early days of the Academy's history, funds for carrying on the operations of the Academy were inadequate; and Bache was one of those who contributed each year from his personal resources. The desire to help others was a part of his nature, and, having in mind the conditions under which scientific research of the day was carried on, he bequeathed to the Academy its first grants-in-aid fund. It accrued to the Academy after the death of his wife in 1871. The capital sum, which originally amounted to about $50,000.00, now totals $60,000.00. During the sixty-nine years which have since elapsed, three hundred eighty-five grants, for the prosecution of researches in physical and natural science, have been made from the income. Added together, the grants amount to $171,762.67; and the average amount of each grant is about $450.00. Thus Bache's efforts on behalf of science have prospered and will, it seems likely, continue to prove of meritorious service well into the future.

TERRESTRIAL ELECTRICITY IN RELATION TO GEOMAGNETISM

O. H. GISH

Assistant Director, Department of Terrestrial Magnetism, Carnegie Institution of Washington

(*Read February 14, 1941, in Symposium on Geomagnetism*)

In this meeting for the commemoration of the life and work of Alexander Dallas Bache who, in addition to many other notable contributions, gave important stimulus to the scientific investigation of geomagnetism in the United States, it seems appropriate at this point also to salute another great American for initiating the first scientific investigation of terrestrial electricity. Nearly 200 years ago, in June 1752, in this city, Franklin made his famous experiment with the kite in a thunder-cloud. The beauty and finesse of this experiment are impressive. Although Dalibard in France had succeeded about one month earlier than Franklin in collecting electricity from the air during a thunder-storm by using an iron rod one inch in diameter and 40 feet long, supported vertically and insulated from ground, he gave Franklin credit for suggesting the experiment when a few days later he enthusiastically reported to the French Academy that "Franklin's idea ceases to be a conjecture, here it has become a reality."

Thus began the first epoch in the development of the science of atmospheric electricity, that branch of terrestrial electricity which deals with the electrical properties and electrical phenomena of the atmosphere.

The first epoch of another branch of terrestrial electricity began soon after the telegraph came into use when it was found that sometimes electric currents, other than those used for sending signals, transgressed upon the wires. In 1847 W. H. Barlow, who was in charge of telegraph lines in England, observed that these transgressing currents appeared on all the lines at about the same time and he concluded that the Earth is the common source of these currents. This branch of terrestrial electricity, which may be called geoelectricity, deals with the electrical properties of the

Earth and the electrical phenomena which transpire within the Earth.

Manifestations, like these, of the natural electric currents in the Earth or conspicuous manifestations of electric forces at play in the atmosphere, such as lightning, St. Elmo's fire, etc., are of relatively infrequent occurrence, but electric forces are always at play in the atmosphere and in the Earth. However, these are usually so feeble that they can be "observed" only with the aid of suitable instruments and equipment.

There is a parallel in geomagnetism. Although the magnetic force at a given place seldom changes enough during short time-intervals to be noticeable with an ordinary compass, yet this force is continuously changing by small amounts. Observations of these inconspicuous phenomena of atmospheric electricity, geoelectricity, and geomagnetism reveal interesting relations between some phenomena in these three categories—relations which aid in an understanding of certain effects in practical affairs which occasionally come to public attention. The occasions to which I refer are usually called magnetic storms. Electromagnetic storms would be a better, although also an inadequate, designation.

A few selections from the case history of one of the most outstanding of these storms will provide illustrations of some relations between terrestrial electricity and geomagnetism. This storm began everywhere on the Earth at about 14 hours after midnight at Greenwich (9 A.M. Eastern standard time), March 24, 1940. During the latter part of that day there was much interference with wired communication-services, and with radio broadcasts, unprecedented obstruction of trans-Atlantic radio communication was reported; and most surprisingly, there occurred some interference with operations of a number of electric power-systems in the United States and Canada.[1]

Magnetic records such as that obtained at the Carnegie Institution's magnetic observatory near Huancayo, Peru (upper left of Fig. 1) showed that this electromagnetic "storm" began everywhere shortly before 14 hours after Greenwich midnight and continued with remarkable activity for about 24 hours.

The activity of the electric currents in the Earth for the same period at Huancayo is illustrated by the graph in the lower left

---

[1] For a description of effects of this electromagnetic storm on (*a*) electric power-systems and (*b*) on communication-systems, see (*a*) W. F. Davidson, *Edison Elect. Inst. Bull.*, **8**, 365–366, 374 (1940), and (*b*) L. W. Germaine, *ibid.*, **8**, 367–368 (1940).

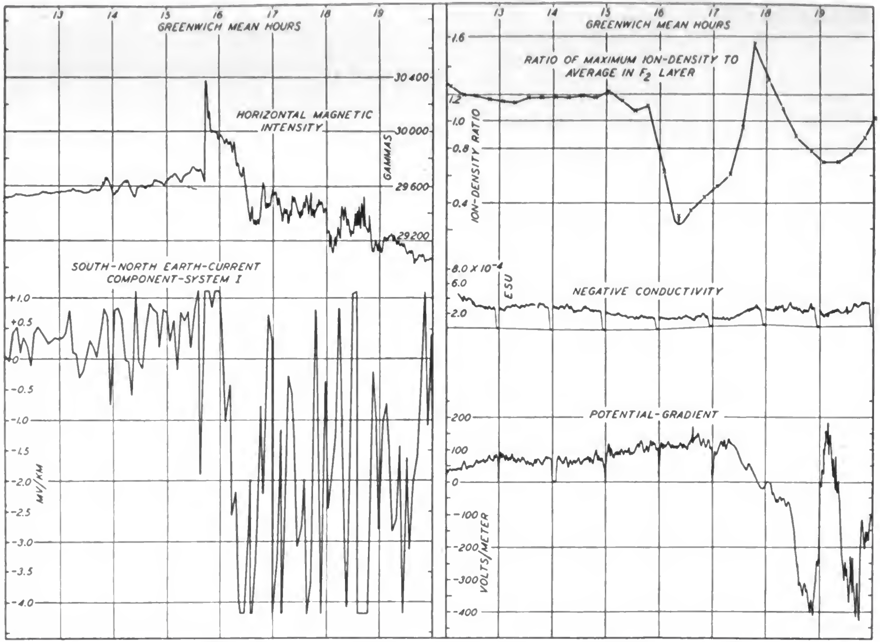


FIG. 1. Terrestrial electric and magnetic records obtained at Huancayo Magnetic Observatory during part of magnetic storm of March 24–25, 1940.

of Fig. 1, where a component (northward acting) of the electric force in the Earth is charted. Certain points of correspondence between these two elements are readily seen. A comparison of the details of the electric record with those in the magnetic record is not justified because the former is an imperfect representation of the earth-current. The recording instrument was too sensitive for the present purpose; hence at intervals during this storm the values were greater than could be recorded, for example, during the interval extending from before until after the time when the magnetic force was greatest. It was in this flood stage of the Earth's electric currents that serious interference with the operation of telegraph-systems and electric power-systems was experienced. This record furthermore does not show the more rapid fluctuations. A more responsive recorder, in use at the Tucson Magnetic Observatory, recorded many rapid oscillations, not shown here, while some insensitive recorders which were put into service by the Western Union Telegraph Company and others during part of the storm indicated that the earth-current at times was probably more than one thousand times the normal intensity. For example, 800 volts difference of potential was measured between the ground-contact at New York, New York, and that at Binghamton, New York, which are 140 miles apart, thus indicating a gradient of 5.7 volts per mile. A somewhat greater intensity (10 volts per mile) was estimated by W. F. Davidson, as necessary to cause the interference observed on power-transmission lines. A greater earth-current intensity than this has been reported only once, namely, between New York, New York, and Elizabeth, New Jersey, on July 16, 1892, when an isolated measurement indicated an average potential-gradient of about 14 volts per mile between New York and Elizabeth.

Smaller effects of this sort are well known to telegraph operators and, I have been told that, sometimes they allude to these as "an aurora on the line." Indeed at such times auroras are likely to be observed at places far from the zone of maximum auroral frequency, but the auroras are certainly not directly related to the disturbance of the telegraph.

Although the records obtained during this storm illustrate a correspondence generally found between the disturbances of the electric currents in the Earth and the geomagnetic disturbances at the Earth's surface, they also show that there is no conspicuous correspondence between the magnetic disturbances and disturbances of the electric force, or of the electric properties, in the atmosphere near the Earth. The lowermost graph in the right-hand side of Fig. 1 represents the vertical electric force, or potential-gradient, in the atmosphere near the Earth for the same period as that for the graphs which have already been examined. Up until about $17^h.5$ there are no prominent departures from the usual type of such records even though the electromagnetic "storm" had been in progress about 3.5 hours, during the last two of which it was in its most intense stage. The disturbance of potential-gradient shown in the latter part of that graph ($17^h.5$ to 20^h) is the result of a small local thunder-storm, such as occur so frequently in the vicinity of the Huancayo Magnetic Observatory, and hence this portion cannot be compared with the magnetic and earth-current records. That thunder-storms are not closely related to these electromagnetic storms is evident from the well-known and fortunate circumstance that thunder-storms do not occur simultaneously everywhere on the Earth.

The graph just above that for potential-gradient is the record of the electrical conductivity of the air near the Earth's surface—that part which depends upon the presence of negative ions, hence the term "negative conductivity." A corresponding record of "positive" conductivity—that which depends on positive ions—was also obtained but since the two are of similar character the one shown here is adequate to illustrate the fact that no conspicuous effect in the electrical conductivity of the air near the Earth is manifested during one of these electromagnetic storms. Furthermore, since the value of potential multiplied by the value of the total air-conductivity (positive conductivity plus negative conductivity) gives a value of the electric current which flows from the atmosphere to the Earth (or at times in the opposite direction) it follows that the electric current from air to Earth also suffers no conspicuous change during electromagnetic disturbances. Such observations constitute one reason for inferring that the electric phenomena and electric properties of the lower atmosphere, probably up to an altitude of more than 20 km, are not involved to an important extent in these electromagnetic disturbances.

However, it was suspected long ago that electrical properties and forces somewhere up in the atmosphere are fundamentally involved in some of the magnetic variations. Balfour Stewart in 1882 concluded that the diurnal variation in geomagnetism should be attributed to systems of electric current located somewhere in the atmosphere. He recognized that this required that the electrical conductivity of the air there be quite large and, although practically nothing was known about the electrical conductivity of air at that time, he thought that the phenomena of polar lights and certain experiments made in the laboratory showed that air at the low pressure which prevails in the high atmosphere may become a good conductor of electricity. Nevertheless this was an obstacle in the way of his theory. But the need for a theory of this sort became even more urgent when Sir Arthur Schuster, six years later, by quantitative analysis, corroborated Stewart's conclusion, that the primary origin of the diurnal magnetic variations is not within the Earth. A few years later the introduction, by J. J. Thomson and his collaborators, of the concept of ions in gases, opened the way for the development of the present understanding of the conduction of electricity in air and other gases. With that the postulated high conductivity in the upper atmosphere seemed more plausible.

Further evidence that the electrical conductivity of air in the higher atmosphere is remarkably large was provided by early experience with radio telegraphy. Professor Kennelly of Harvard and Oliver Heaviside, British engineer, independently concluded in 1902 that, in order to account for the fact that radio messages were received at unexpectedly great distances from the sending station, it was necessary to assume that the conductivity of some region of the high atmosphere was so great that radio waves are reflected back toward the Earth and thus reach beyond what would otherwise be the radio horizon. Professor Kennelly estimated the altitude of this reflecting surface to be about 80 km, which is close to later measurements.

In recent years methods and equipment have been developed, and are now in use at a number of places, for obtaining automatic records which provide a measure of the effective altitudes of the highly conducting strata—for now it is known that there are several of these at different altitudes—and of the maximum concentration of ions in these strata. From such measurements it appears that the concentration of ions, and the corresponding conductivity, is great enough to admit of a theory which has been developed more or less along the lines suggested by Balfour Stewart.

The character of the variation of the maximum ion-concentration, or ion-density, of the high atmosphere in the uppermost of these strata, during part of the electromagnetic storm of March 24, is represented by the graph in the upper right-hand part of Fig. 1. The ordinate there represents the ratio of the maximum ion-density on the disturbed day to the maximum ion-density at the corresponding time on an undisturbed day. The magnitude of the changes during the more intense part of the storm is impressive. Interest in this is enhanced by the fact that it is a record of a result produced at a very high altitude in the atmosphere. The altitude of this layer, on a normal day at Huancayo, ranges between 300 and 400 km but as Mr. Berkner, who kindly provided the data for this graph, will show in his address (page 309), it varied considerably during this storm.

Another electrical aspect of the atmosphere which was registered simultaneously with all the other elements exhibited here, namely, the intensity of the cosmic radiation, varied in the manner shown in Fig. 2. Beginning early on March 23 and extending into March 26 the percentage-departure from the mean intensity of the cosmic radiation is there charted. At about the beginning of the storm (14ʰ GMT) the intensity fell below normal and continued with diminished intensity throughout the following day. The greatest diminution is about three per cent. This is an interesting relation with geomagnetism which may be passed without further comment because it will doubtless be discussed in the next session of the program.

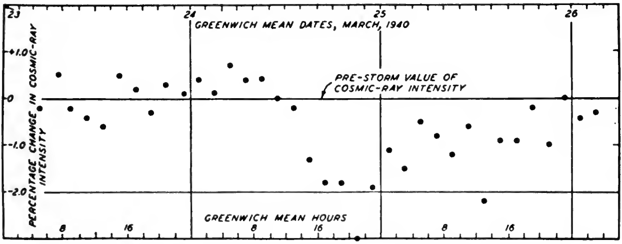


FIG. 2. Cosmic radiation, Huancayo Magnetic Observatory, during magnetic storm of March 23–25, 1940.

In this brief outline of the case history of the electromagnetic storm of March 24–25, 1940, some relations between terrestrial-electric phenomena and properties and phenomena in geomagnetism have been cited or illustrated. The association of polar lights, or auroras, with electromagnetic storms, the changes in the electric properties of the high atmosphere (ionosphere) at such times, which involve deterioration of the mirror-like reflection of radio waves from the ionosphere back to Earth, and the diminution of the intensity of cosmic radiation, are features which will be discussed by other speakers. The rest of this discussion therefore may be restricted to terrestrial-electric phenomena and properties, of the lower atmosphere and of the Earth, in relation to geomagnetism.

The fact that no conspicuous changes in the electric force or in the electric conductivity of the lower atmosphere have been observed during electromagnetic storms has already been mentioned. The question whether there may be some minor changes is not yet settled. A few investigators have reported what they regard as evidence of such effects but there are important conflicts between the several lines of evidence. Apparently, any such effects will be definitely revealed only by more elaborate statistical analyses. Despite the fact that no positive relationships between the terrestrial-electric phenomena of the lower atmosphere and geomagnetism have been established, knowledge of the former has played an important rôle in the development—or more correctly, the restriction—of explanations of the phenomena of geomagnetism. Thus several explanations which have been proposed to explain the general magnetic field of the Earth, or to explain special aspects of that field, have had to be abandoned because they implied the existence of vertical electric currents in the lower atmosphere of a magnitude many many times that found from atmospheric-electric measurements.

In contrast to the electric phenomena in the lower atmosphere, the electric currents within the Earth, as already indicated, bear a close relation to temporal changes in geomagnetism. The "stormy" periods for the earth-currents correspond closely with those of geomagnetism. This as well as the fact that many of these storms are world-wide in scope and run a similar course everywhere simultaneously, is illustrated by Figs. 3 and 4.

The graphs in these figures are copies of records obtained at

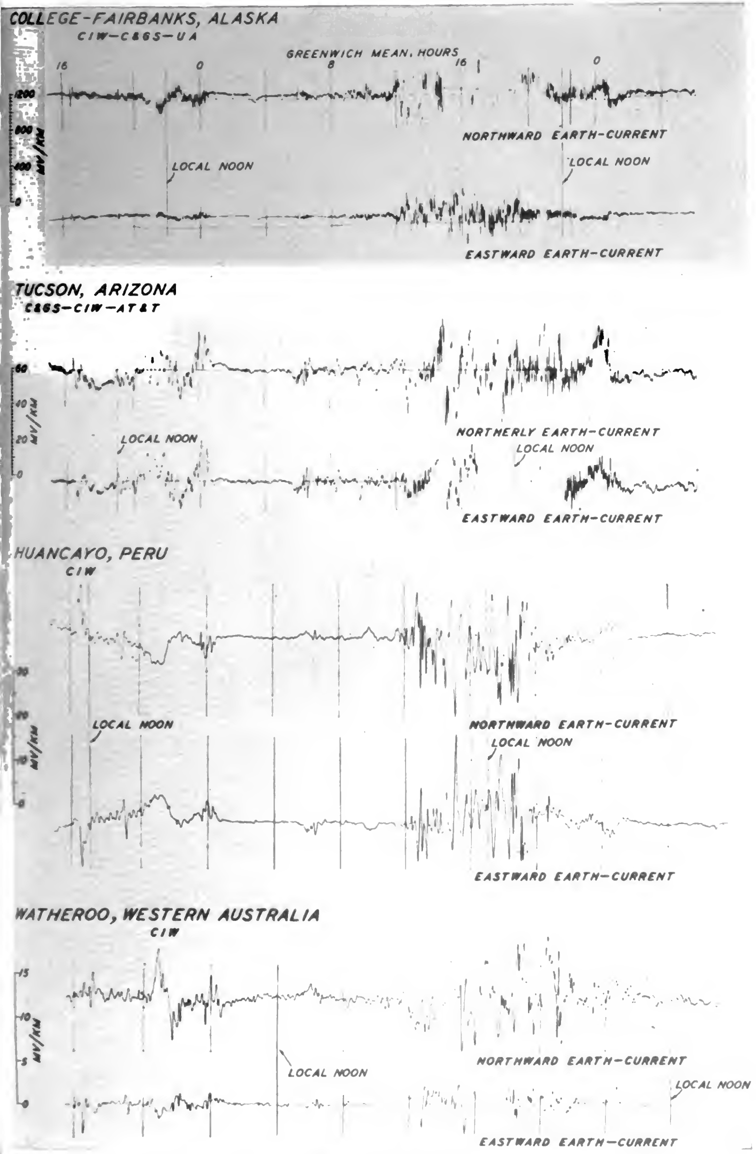


Fig. 3. Earth-current records, storm April 30 to May 2, 1933.

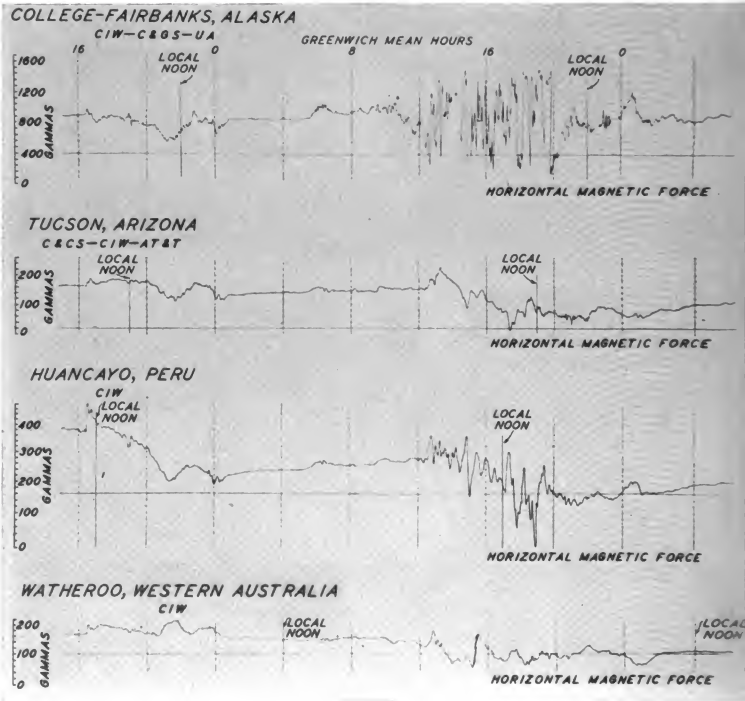


Fig. 4. Geomagnetic records, storm April 30 to May 2, 1933.

four places ranging in latitude from 65° north to 30° south during 15^h GMT, April 30, to 06^h GMT, May 2, 1933. In Fig. 3 two components of the horizontal electric potential-gradient in the Earth are shown for each place. These are designated as "northward earth-current" and "eastward earth-current." The significance of these terms may be brought out by the following example: If the recorded value of the components is greater than the mean (above the mean position on the graph) at a given instant then the current in the Earth, corresponding to the measured gradient components, flows in a northeasterly direction, and if the northward component is positive (above the mean position) while the eastward component is negative (below the mean position) the current flows in a northwesterly direction. The amplitude of deflections from the mean are in a fixed relation to the intensity of the current at a given place but the amplitudes for one place cannot be compared with those for another without taking into account the electrical conductivity of the Earth. It may be mentioned here that the direction of the current in the Earth doubtless varies somewhat from normal from place to place on account of the irregular distribution of the electrical conductivity of the Earth and that the distribution of electrical conductivity to depths of many miles in the Earth is a factor of importance for some aspects of geomagnetism. The complexity introduced by irregularities in the distribution of conductivity does not affect the times of occurrence of the disturbances in earth-currents. As is to be seen from Fig. 3, the disturbances at different places correspond quite closely in time. They also correspond well with the periods of magnetic disturbance shown in Fig. 4 by the reproductions of records of the horizontal magnetic intensity.

These extraordinary world-wide electric and magnetic impulses, which occur irrespective of daylight or darkness, in fair or foul weather, are however apparently related to activity on the Sun because it is found that they generally are more frequent in years when sunspots are more numerous. Another indication of the relation to solar activity is the fact that these electromagnetic impulses tend to recur after an interval of about 27 days—close to the period of the Sun's rotation and definitely less than a lunar revolution. Other less intense electromagnetic impulses also wax and wane with the sunspot-cycle. Fig. 5 is an illustration of the latter. The polar lights also appear to be dependent upon solar activity in about the same way as these several electromagnetic phenomena, and evidence accumulated in the last two decades from effects in radio transmission, or from special radio experiments, appears to show that the electrical conductivity of the high atmosphere depends, in a similar way, upon solar activity.

In contrast to the impulsive electromagnetic phenomena which have just been reviewed there is a more regular tranquil aspect. Both the electric gradient in the Earth and the magnetic force vary in a fairly regular manner during the day (diurnal variation). Although some impulsive movements occur at irregular intervals nearly every day yet on so-called "quiet" days, these are so feeble

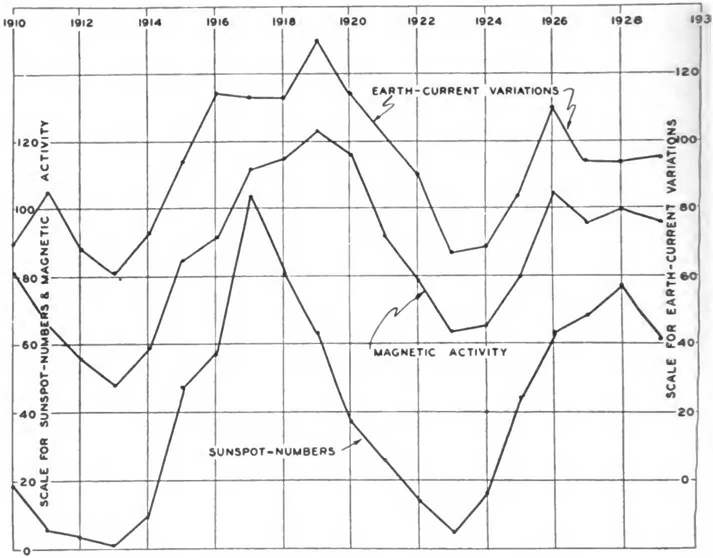


Fig. 5. Dependence of electromagnetic variations upon solar activity.

that the gross features of the regular diurnal variation can be seen in the record for an individual day. Such is the case for the records of the two earth-current components registered at the Tucson Magnetic Observatory of the United States Coast and Geodetic Survey, on January 9 to 11, 1933, which are exhibited in Fig. 6. At night the variation is small but in the forenoon the electric gradient in the Earth increases to a maximum at which time the earth-current at Tucson flows in a northeasterly direction. Following this the gradient decreases and shortly before noon reverses direction until shortly after noon a maximum current-intensity is again reached but this is directed roughly toward the southwest. There is another cycle much like this but the maximum current-intensity is less. If the average ordinate for each hour is taken from such records and the mean of those for each of the 24 hours of the day is obtained for a period of a month or a year and charted, the graph is more regular, and is more suitable for quantitative study.

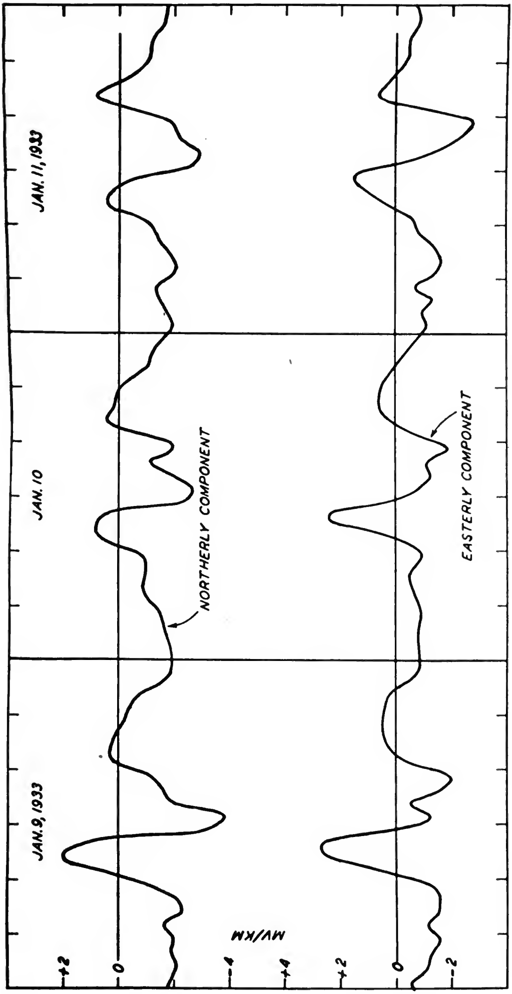


FIG. 6. Hourly values of northerly and easterly earth-current potential-gradient in three successive (January) calm days at Tucson, Arizona.

The average character of the diurnal variation, for each of three different places [Berlin (Germany), Ebro Observatory (Tortosa, Spain), and the Watheroo Magnetic Observatory (Western Australia)], is shown in Fig. 7 by a full-line graph derived in the manner just indicated. The broken-line graphs in Fig. 7 were calculated with the aid of formulæ developed by S. Chapman and T. T. Whitehead. These formulæ are essentially the quantitative expression of a corollary of the Balfour Stewart theory for the diurnal variation in geomagnetism as developed in quantitative form by Schuster and amplified by Chapman. Considering the fact that in the development of these formulæ no account was taken of those irregularities in the structure of the Earth's crust which present great contrasts in electrical conductivity and thus certainly distort the electric flow, it seems that the agreement of the graphs for the calculated values with those for the observed values is good enough to justify the conclusion that in the main the diurnal variation in the earth-current and that in geomagnetism are two electromagnetic manifestations of electric current-systems in the high atmosphere.

Thus far in speaking of the diurnal variation of both earth-currents and geomagnetism no mention has been made of an interesting influence of the Moon on these phenomena. The latter is eliminated when mean values of the measured elements are formed, for each hour of the solar day, over a long period, say a year or more. The result then represents only the average solar diurnal variation. But if mean values are formed in a manner similar to the foregoing for the respective hours of the lunar day then in a long series the solar diurnal variation is eliminated and the result represents the average lunar diurnal variation. This lunar component has a much smaller amplitude than the solar component but it is a well-determined component of the diurnal variation of both earth-currents and geomagnetism. This component is attributed to tidal action of the Moon upon the high atmosphere, which moves the highly conducting air back and forth across the magnetic lines of force of the permanent magnetic field of the Earth, and thus an electric current is generated in that region. Similarly winds of other than tidal origin doubtless play a rôle in generating the electric currents to which the solar diurnal variations may be attributed. Fig. 8 is designed to convey a conception of the electric current-system in the high atmosphere to

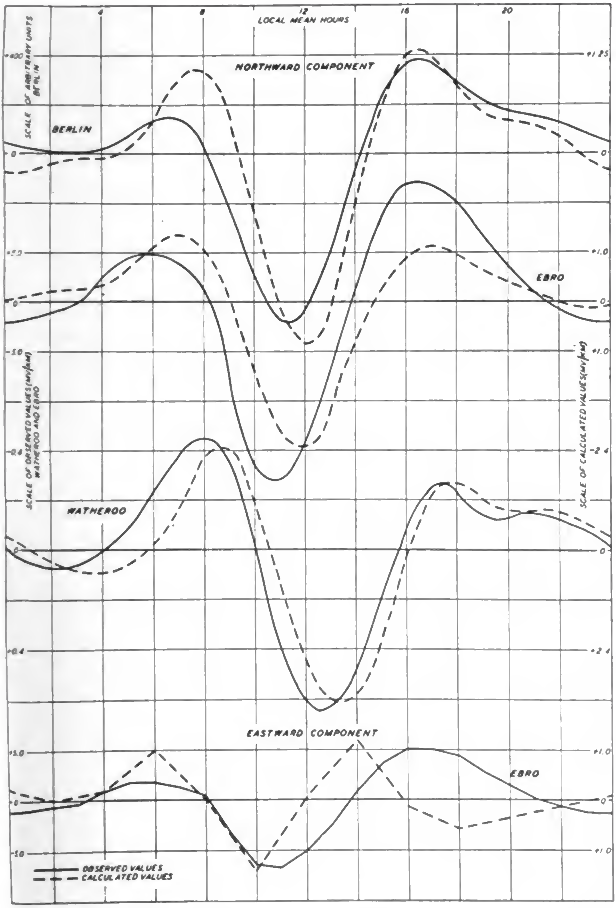


FIG. 7. Comparison of theoretical and observed diurnal variation of earth-currents.

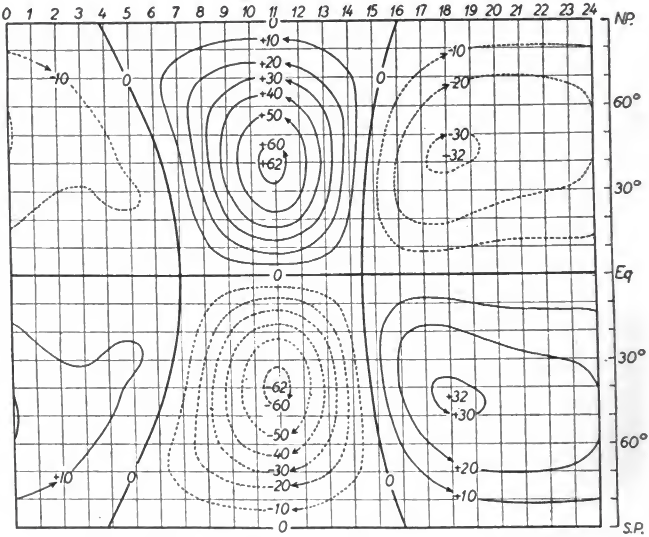


FIG. 8. Electrical circulation in the high atmosphere deduced from magnetic diurnal variation (after Bartels).

which the solar diurnal variation in geomagnetism is attributed. In fact, it is derived from the latter.

Fig. 8 is a map of the current-system at the time of equinox as viewed from the direction of the Sun. The position at the center of the map is always directly under the Sun. Between two adjacent lines 10,000 amperes of current flow in the direction of the arrows. The large whirl of current in the upper center lies over the Northern Hemisphere and gives rise to a magnetic field of opposite sign to the corresponding whirl in the lower part which lies over the Southern Hemisphere. In both cases the vertical component of this field opposes the vertical component of the Earth's permanent field, the diminution being greatest at places which happen to be under the centers of these whirls. The Earth rotates within the system; hence, a place which at about 11 A.M. was under the center of the principal whirl *A* of the Northern Hemisphere would be under the center of the minor whirl *B* at

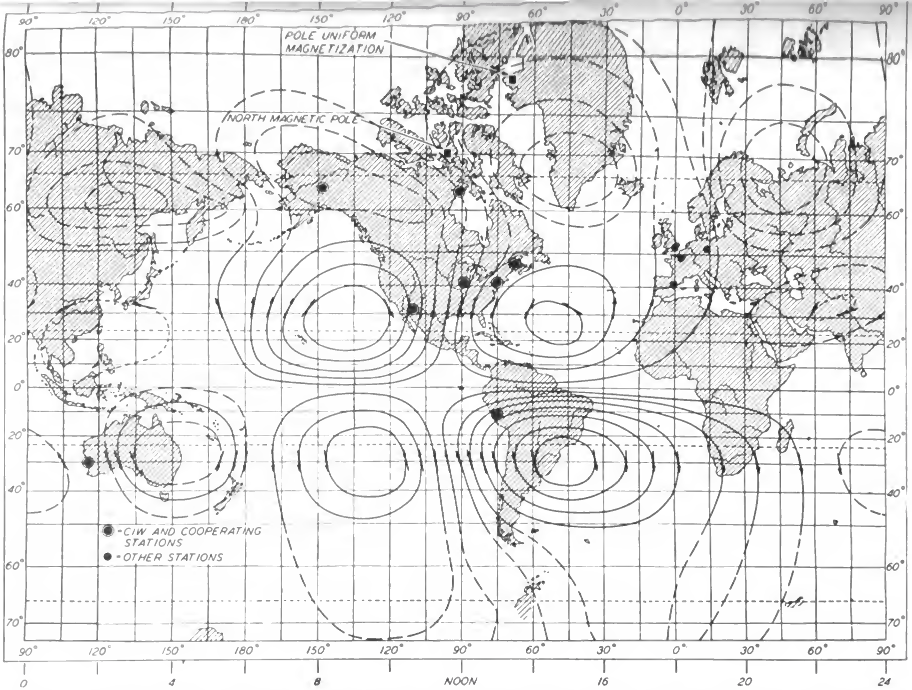


Fig. 9. Equatorial view of position of earth-current system at 18^h Greenwich mean time as deduced from observations.

about 6 P.M., and since the current here circulates in the opposite direction an observer would then observe an increase in the vertical magnetic force. Thus as the Earth rotates while this system remains fixed, in respect to the Sun, geomagnetic variations which follow a regular daily schedule occur everywhere. However, the rotating Earth and the fixed current-system provide all the elements for a dynamo, in which the current-systems, A, A', B, and B', provide the poles, and the Earth serves as the armature. Obviously then currents must be generated in the Earth. These earth-currents in turn give rise to a magnetic field at the Earth's surface which, as is found from analyses of data for the diurnal variation of geomagnetism, has an average strength of about one-third that of the currents in the high atmosphere. The net diurnal magnetic variations are therefore a composite, part arising from the external current-system and part from the internal current-system (earth-currents).

A map of this internal current-system, based on measurements of the horizontal electric gradient in the Earth, is shown in Fig. 9. Here the lines of current-flow have a similar significance to those in Fig. 8, but, while these are drawn so that the amount of current which flows between any adjacent pair of solid lines is everywhere the same, that amount is not known. This map serves chiefly to show the position of the several electric whirls and the relative intensity of current and its direction of flow. Of course, the actual current-system is more complicated than is indicated here, chiefly because of the very irregular distribution in the Earth of good and bad conductors of electricity. However, this is thought to represent approximately the more prominent features of the electric current-system in the Earth on so-called quiet days, and presents an integrated view of that aspect of terrestrial electricity in which investigation of the relations to geomagnetism have made the greatest advance.

THE MAGNETIC SURVEY OF THE UNITED STATES

N. H. HECK

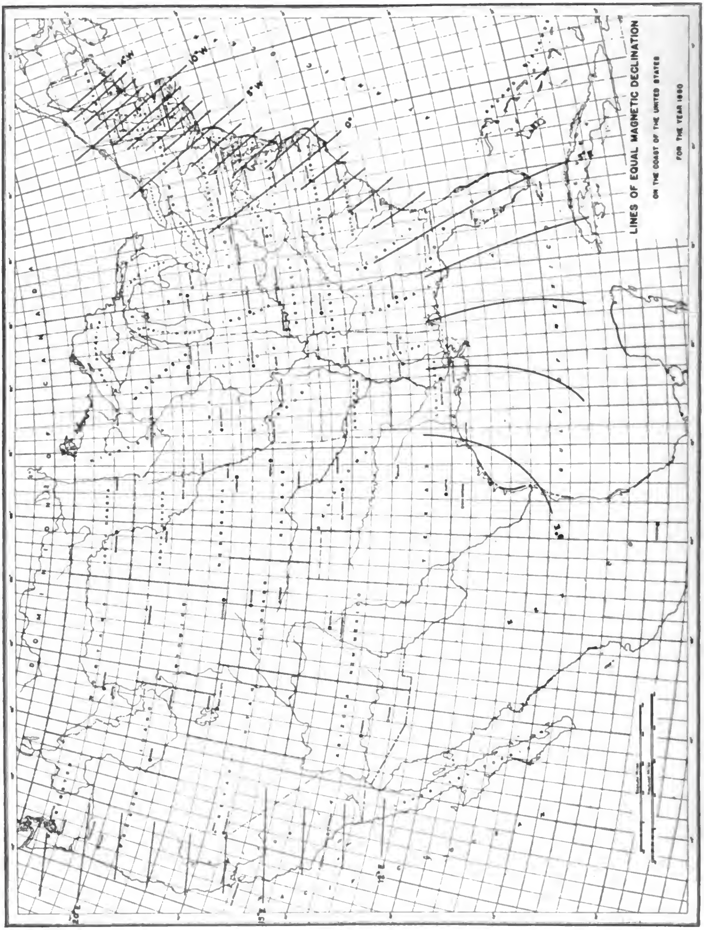
Chief, Division of Geomagnetism and Seismology, United States
Coast and Geodetic Survey

(*Read February 14, 1941, in Symposium on Geomagnetism*)

Many of the papers in this symposium deal with outgrowths of the developments initiated in Philadelphia 100 years ago by Alexander Dallas Bache, which either were not recognized to be associated with the earth's magnetism or were not then dreamed of. The magnetic survey of the United States, on the other hand, was practically started by Bache and there has been no break in its continuity.

A limited number of magnetic observations had been made with reasonable accuracy in connection with land surveys and a few for the purpose of the magnetic survey had been made when he became superintendent of the Coast Survey in December 1843. Shortly after, in addition to his scientific interest in this work, he recognized the practical importance in connection with navigation.

It is interesting to note that the impulse to development of scientific work in both astronomy and geomagnetism came from the same source—the practical needs of navigation. It is to Bache's credit that he went beyond the immediate pressing needs and at least began the scientific attack on the problems of the magnetic survey of the United States. By 1855 it was possible to prepare lines of equal declination or isogonic lines, for the coasts of the United States for 1850 (Fig. 1) and within a few years after his death, the 1870 map showed isogonic lines for all of the country except the western plains and part of the Rocky Mountains. The lines were much smoothed and uncertain, but the basis for present-day practice was there. By 1875 the lines were given for the whole country (Fig. 2) and in 1885 maps appeared for all the elements, but with still great unevenness of distribution of stations. In 1900 maps were again prepared for all the elements but many lines were dotted as uncertain and approximate rates of change were shown. By 1902 the present practice regarding maps was


LINES OF EQUAL MAGNETIC DECLINATION
ON THE COAST OF THE UNITED STATES
FOR THE YEAR 1850
14°W
10°W
5°W
0°
5°E
15°E
20°E
DOMINION OF CANADA

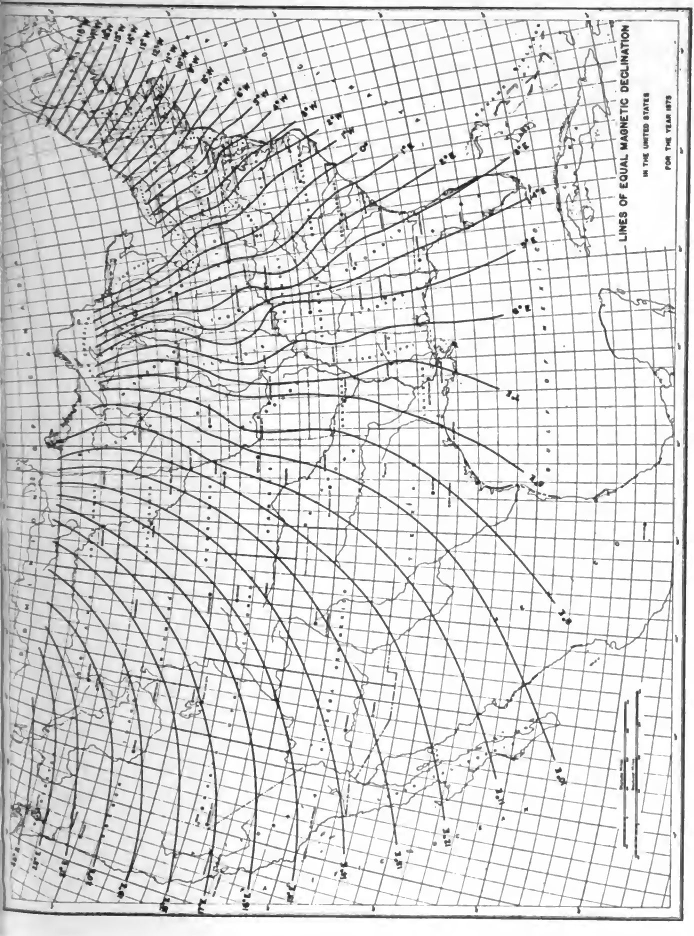

LINES OF EQUAL MAGNETIC DECLINATION
IN THE UNITED STATES
FOR THE YEAR 1875

Fig. 2.

adopted, and the rapid increase in observations began to give significance to the maps. This is illustrated by the most recently published map, that for 1935 (Fig. 3).

While much of the development after Bache's death is to be credited to C. A. Schott, L. A. Bauer, R. L. Faris, D. L. Hazard and others, it is true that Bache started the work with vision and with foresight of problems which time did not permit him to solve and some of which indeed are not yet fully solved.

Bache not only spent much time studying the Girard College results but was responsible for the installation of a magnetic observatory at Key West, Florida, in 1860, where we would like to have one today. The value of his broad viewpoint is evident from the success of the plan proposed and partly executed by L. A. Bauer under the general direction of Henry F. Pritchett for five observatories, all with similar equipment, and the occupation of magnetic stations, at least one in every county, preferably at the county seat. All this provided essential preparation for the work of Bauer and Fleming in developing the world-wide attack on the problems of geomagnetism by the Department of Terrestrial Magnetism of the Carnegie Institution of Washington. There has ever since been close cooperation between the two organizations and certain essential public services could not have been provided without it.

We have then the picture of a few observations near the coasts to meet the needs for information for the nautical chart with gradual extension into the interior, aided by special surveys, such as those by L. A. Bauer in Maryland, F. E. Nipher in Missouri, and by the State Geological Survey of North Carolina in cooperation with the Coast and Geodetic Survey. Then in 1899 came the plan for extending the survey to all parts of the country and territories and island possessions. During the period from 1860–1899, there was at one place in the United States a magnetic observatory in operation though the place was shifted a number of times (see p. 239).[1] The extended plan provided for 2 observatories in the United States proper—Maryland (Cheltenham near D. C.), and Kansas (Baldwin)—one each in Puerto Rico (Vieques), Alaska (Sitka), and Hawaii (near Ewa, Oahu). In 1909 an observatory in Arizona replaced that in Kansas, and in 1926 the Puerto Rico observatory was moved to the vicinity of San Juan.

---

[1] H. E. McComb, Geomagnetic Observatories and Instruments, *Proc. Amer. Philos. Soc.*, **84**: 239–255 (1941).

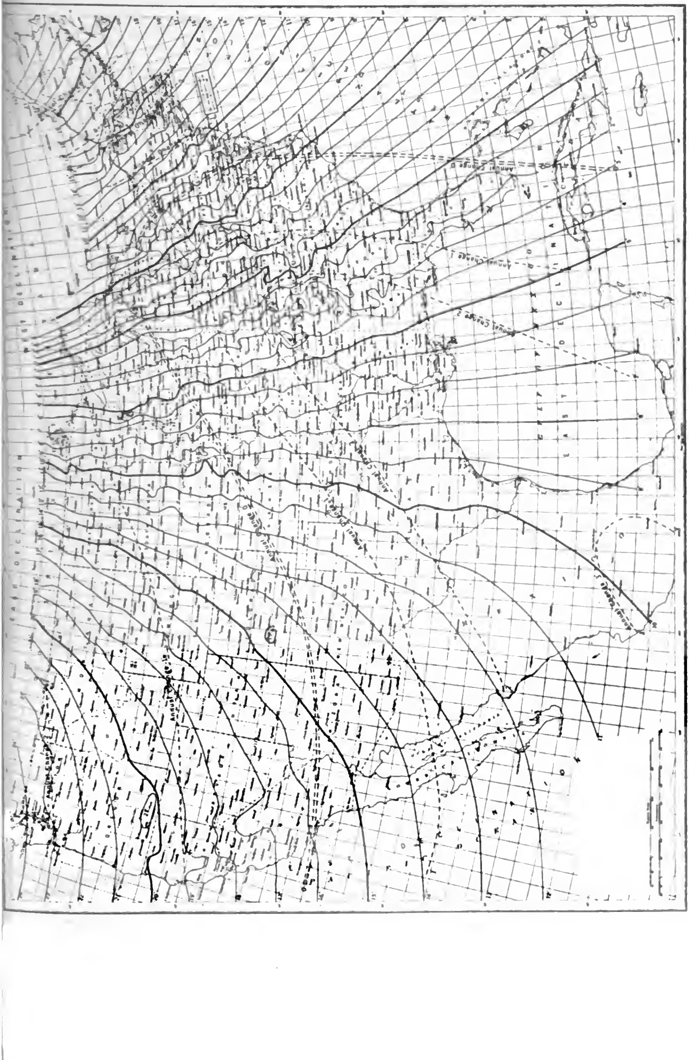


FIG. 3. Lines of equal magnetic declination and of equal annual change in United States for 1935.

The survey of the entire United States has been a major undertaking. The occupation and maintenance of enough stations for an area of 3,000,000 square miles of the continental United States, not to mention Alaska, Hawaiian Islands, Philippine Islands, and smaller island groups, has required thousands of observations. For the United States proper, most of the expansion was between 1899 and 1918. Since that time, principal stress has been laid on maintenance.

The purpose of the magnetic survey of the United States is to know for any given period what is the magnetic distribution; to find the secular change over a given period—usually 5 years; to establish new stations; to replace those which have been destroyed or have become unsuitable for magnetic observations; and to survey regions of magnetic anomaly. The magnetic survey of the United States and of regions under its juridisdiction consists of observations of all the magnetic elements at most stations and declination only at others, and the operation of magnetic observatories at which are obtained continuous photographic records of the various changes in geomagnetism.

Magnetic Stations. A magnetic station consists of a marker, usually a small concrete monument in the top of which a bronze disk is inserted to indicate its purpose and to mark the exact point over which the observation is made. In addition, there must be an object or objects at some distance whose true azimuths are known in order to obtain declination. Observations are made of declination, dip, and horizontal intensity. From these observations, vertical intensity and total intensity can be deduced. A magnetic observatory consists of a group of buildings well provided with observation piers on which are mounted both variation instruments for recording continuously the changes in all elements and absolute instruments with which observations are made at regular intervals so that the continuous observations can be made absolute—that is, expressed in absolute units—and also so that the performance of the various instruments can be checked.

All survey and boundary monuments are subject to loss primarily through extension of civilization, especially construction work of all kinds, and changes due to nature, as floods, earthquakes, and erosion. Many of them can be replaced if lost. The magnetic station is subject to another source of loss—the introduction of magnetic material or of electric direct currents.

Erection of buildings nearby, dumps containing iron or steel, underground pipes, and direct current installation even at a distance if there is leakage, make an otherwise satisfactory station useless, and replacement must be at a new site.

For these reasons the selection of magnetic stations at county seats—while logical in the days of train and wagon travel and when the requirement that county surveyors should yearly test their magnetic compasses was important—was poor from the viewpoint of stations free from artificial magnetic disturbance and their preservation in that condition.

An attempt has been made to use triangulation stations of the widely expanded net in recent years. These stations have the advantage of being located accurately and of having convenient azimuth marks so that the observer does not have to make any astronomical observations. However, though having the advantage of being in the open country, many are placed too near to roads whose passing traffic affects the magnetic field observed.

The most recent idea is to have a limited number of sites chosen with special care both for present and future effects.

The measurement of magnetism more resembles that of the weather than the observations ordinarily made in surveying. The quantity being measured is constantly changing both in a partly systematic and partly irregular way. Some of the irregular changes are large and sudden. The observations at any point have only limited significance if they cannot be coordinated with others and the device is adopted of reducing the observations to a given epoch. In this way only can magnetic maps be made. Further, at times there may be considerable change during the period of observation. For both these purposes continuous observations are needed somewhere in the general region of the field survey. In practice corrections for diurnal variation of the declination are deduced from an observatory, in many cases at a distance, or from more than one observatory. The corrections on quiet days are satisfactory from a single observatory for an area the size of Great Britain, say within a radius of 300 miles. V. Vacquier (1), of the Gulf Research and Development Company, has made an investigation of vertical intensity in the United States and finds that distance at which continuous observations apply, varies with conditions and may be very small. The actual distances between observatories of the Coast and Geodetic Survey range from 1700 to 2700 miles. A general picture of their distribution in conjunction with similar observatories in the Western Hemisphere is depicted in Fig. 4. In interpolating the correction for a given station the value deduced from that at one or more observatories is adopted. While not a satisfactory practice, the

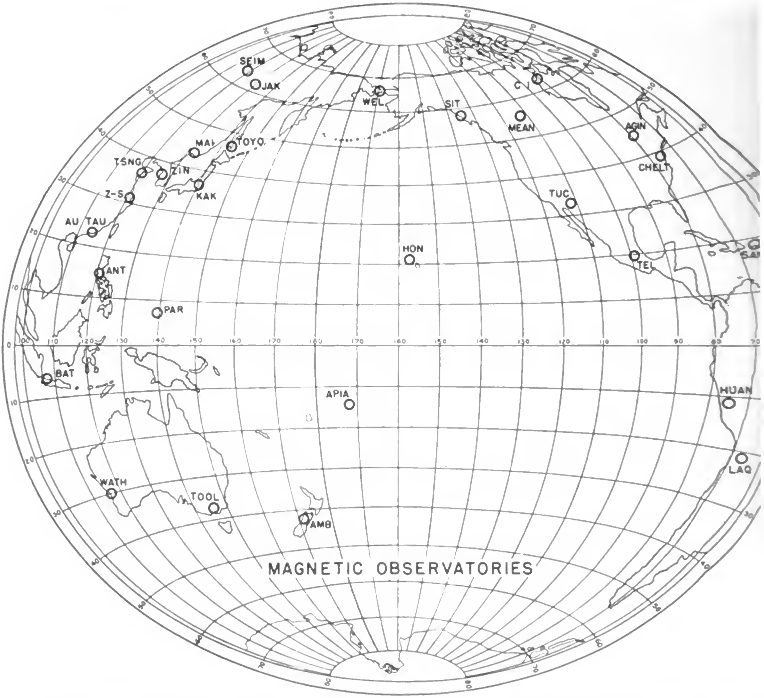


FIG. 4. Showing distribution of coast and geodetic survey magnetic observatories in relation to other

distribution of observatories is fairly good if we include the Canadian observatories at Agincourt near Toronto and at Meanook, north of Edmonton, Alberta, which contribute as much to the work in the United States as they receive from observations at Cheltenham and Sitka. However, the corrections at stations far from these observatories are uncertain. For this reason there is now in the process of development a portable observatory which can be taken to twenty or thirty places in the United States, and operated for several weeks while the survey of the surrounding area is being made, thus furnishing better control for diurnal variation corrections and reductions to mean values.

Absolute determination of vertical intensity. There is a problem in the magnetic survey which had no practical importance until recent years and is yet unsolved except in England and Denmark. This is the measurement of absolute vertical intensity with the same accuracy with which differences in vertical intensity from place to place can be measured. The reasons for this are primarily instrumental. The importance lies in the fact that if satisfactory absolute values were available, a vertical intensity map of the country in considerable detail could be made from various sources such as State Geological Surveys (Missouri, for example) and various commercial organizations might be willing to release information that has lost commercial significance. The point is that these various surveys cannot be combined to form part of a national survey without absolute control.

Recent attempts have been made to solve this problem especially by E. A. Johnson of the Department of Terrestrial Magnetism, Carnegie Institution of Washington, and V. Vacquier, and it is hoped that eventually it will be solved.

It is fortunate that in addition to vertical intensity being the element best suited to give information about magnetic anomalies the disturbance changes in this element in latitudes where most of the work is done are less than for the other elements. This advantage decreases, however, with increased distance from the magnetic equator.

Secular Change. A magnetic survey is not like a geodetic survey. While having definite value in fixing the situation at a given time, it is ephemeral; in fact, there are changes before the map is published. The slow change of average values with time over a period of years is known as secular change. Observatory records show that this change is irregular and the rate of change varies with time. Secular change of magnetic declination at United States observatories is shown in Fig. 5. In practice, repeat observations are made at many stations throughout the country and the best possible adjustment is made to get the rates of change over a five-year period.

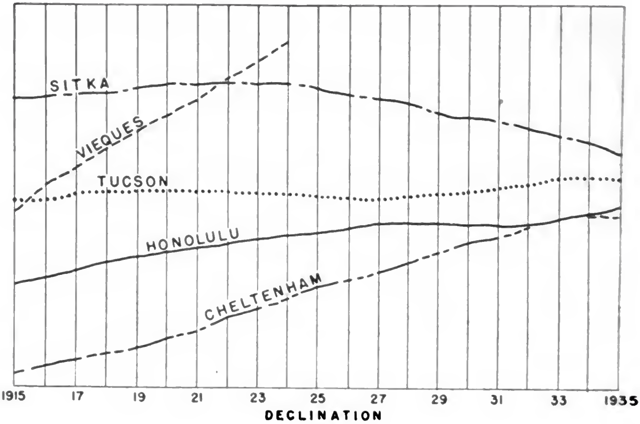


Fig. 5. Secular variation at U. S. observatories.

It is of interest that the standard tables for secular change published by the Coast and Geodetic Survey begin with 1845 or early in Bache's career as superintendent of the Coast Survey.

One of the practical uses of knowledge of secular variation is to derive the values of the magnetic elements at a given place and at any time in the past for which information is available. Many persons, including county surveyors, have made the erroneous assumption of uniform rates of change in one direction. Fig. 6 shows the successive positions of the line of no change at ten-year intervals and that for many years the change was rather uniform—at least it could be fairly well predicted and Schott did so. In 1900 a sudden change occurred and the line instead of being north-south became approximately east-west. Then about 1935 a tendency to form a closed loop or focus began to develop. The results for 1940 just becoming available, indicate that this tendency still persists and is apparently developing.

We do not know the fundamental reason for this change, but it is part of a world-wide phenomenon, as illustrated in Fig. 7. The lines of equal annual change or isopors were prepared by Fiske of the Carnegie Institution of Washington for the period of 1920. The maxima are known as isoporic foci. The positions of the

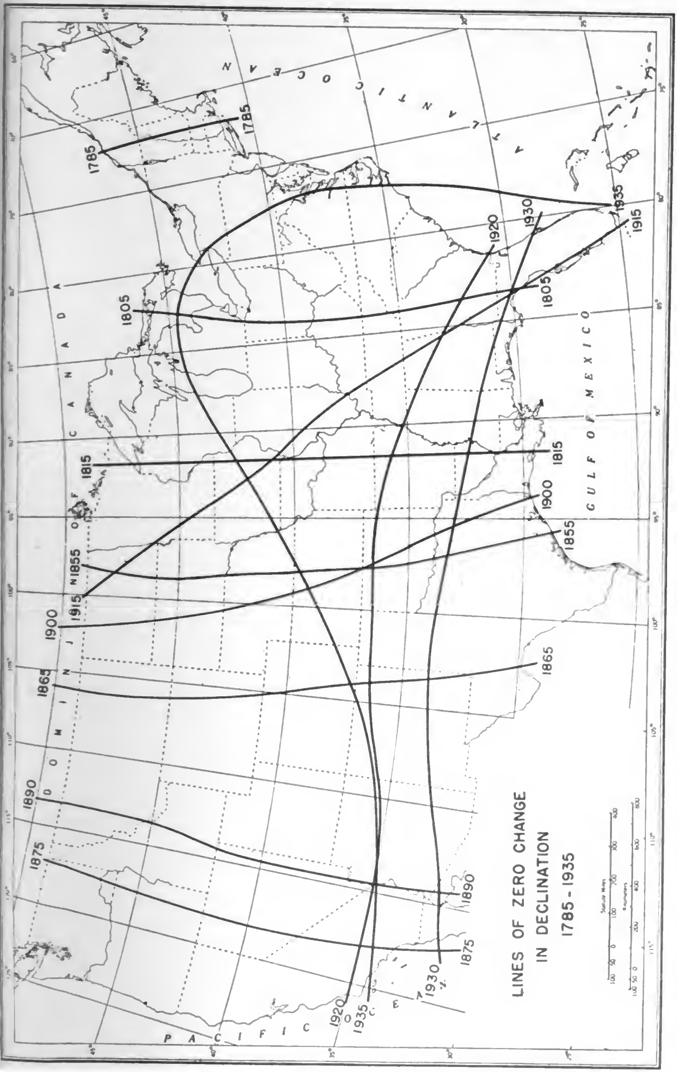

LINES OF ZERO CHANGE
IN DECLINATION
1785 - 1935
ATLANTIC OCEAN
GULF OF MEXICO
PACIFIC OCEAN
DOMINION OF CANADA
1785
1805
1815
1855
1865
1875
1890
1900
1915
1920
1930
1935

FIG. 6.

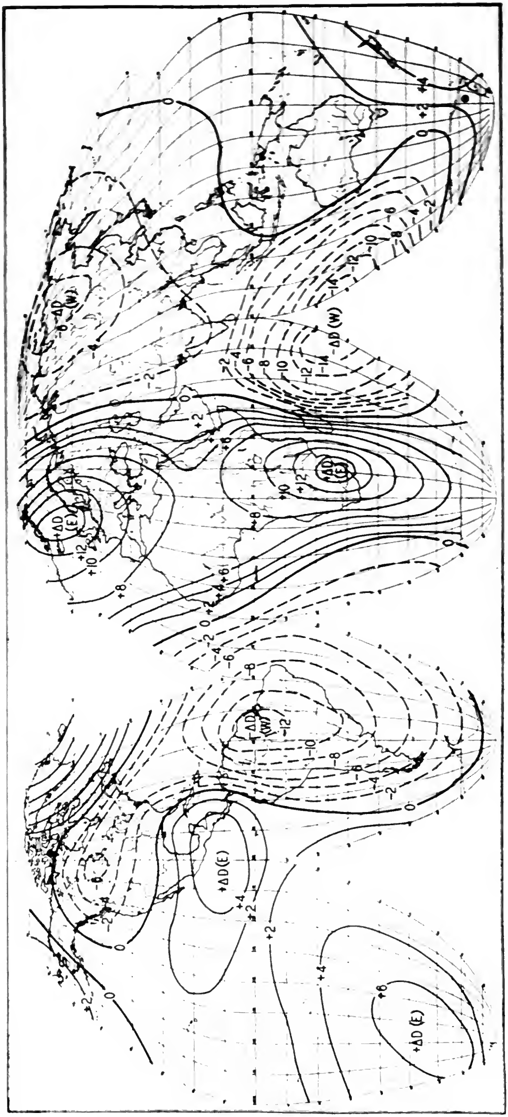


FIG. 7. Isoporic chart for declination (lines of equal annual change) approximate epoch 1920-1925 (After Fiske).

isopors in the United States are influenced by the South American focus. It is shown by the Fiske map and by later data obtained for 1935 for South Africa that both the position and the amount of maximum of such foci change with time.

In spite of certain defects in our program and the more intensive work in Europe, the United States is the largest area over which repeat observations have been made for a period of 40 years with similar methods and standards of accuracy and for five-year periods.

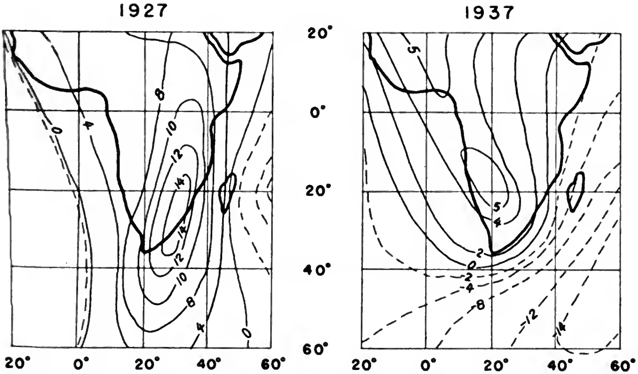


Fig. 8. Isoporic focus in declination at South Africa.

Magnetic Anomalies. If the earth were uniformly magnetized, all iso-magnetic lines would be regular and values could be readily deduced from one place to another. However, the fact that this is not so is both an advantage and a disadvantage. It is this lack of uniformity of magnetization that makes possible the magnetic method of exploration for oil and useful minerals or the determination of broad geological structure. While magnetic deposits on a large or small scale are quite as normal as their absence, such an area is called a magnetic anomaly. The effects may range from detectible small differences to local magnetic poles, as in Siberia and Alaska. Isogonic lines in the vicinity of Port Snettisham, southeast Alaska, are shown in Fig. 9 while Fig. 10 illustrates the Kursk anomaly in Siberia. The anomalies may be highly

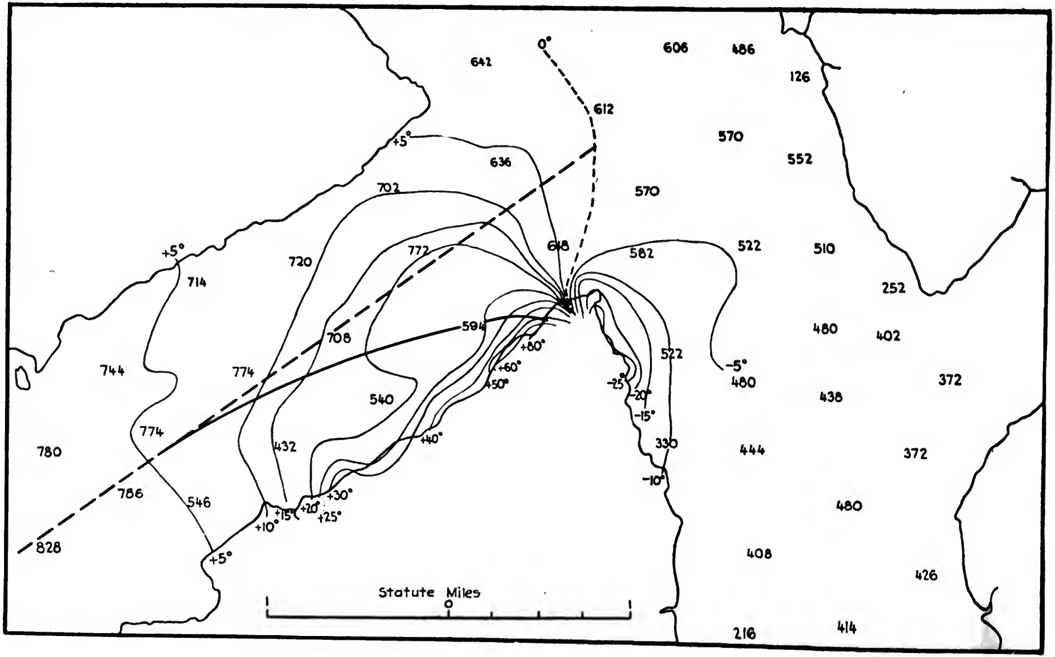
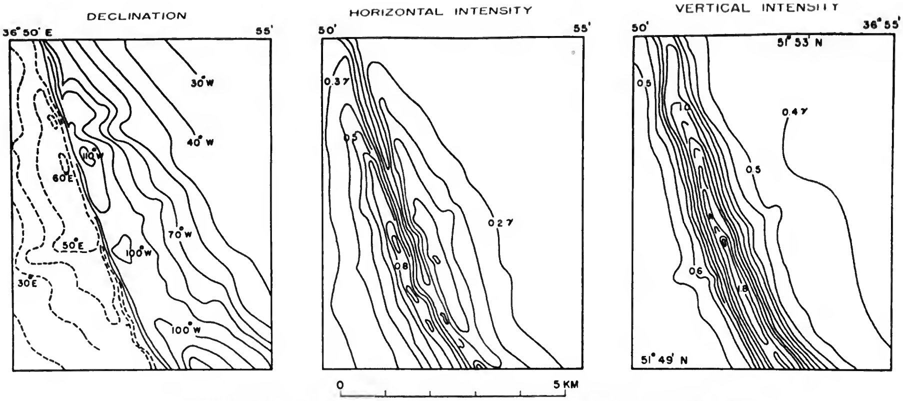
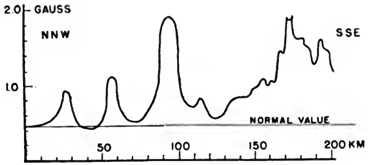


FIG. 9. Magnetic declination anomaly, Port Snettisham, Alaska.



ISOMAGNETIC CHARTS OF THE NORTHERN STRIP OF THE KURSK ANOMALY. (AFTER LASAREFF)



PROFILE FOR THE VERTICAL INTENSITY Z ALONG THE AXIAL LINE OF THE KURSK ANOMALY. (AFTER CHAPMAN & BARTELS)

Fig. 10.

localized or spread over extended areas. Surveys of volcanic areas have been made in Japan and Guatemala.

Though some areas have been thoroughly examined, most areas are partially unexplored. In northern Delaware and just west of Washington, D. C., there are known areas of moderate disturbance. Missouri has large disturbed areas. Areas of extended anomalies because of recent or ancient lava flows on a large scale are included in northwestern United States and many parts of Alaska, including the Aleutian Islands group. Much of the information comes from declination surveys by the Coast and Geodetic Survey and the inference is that there are many undetected areas back from the coast.

In making magnetic observations in the field, special care must be taken to note the height of the instrument above the ground. Chapman and Bartels (2) state that if this precaution is taken, secular change will be found to be the same in disturbed areas as at undisturbed stations in the vicinity. In a highly disturbed region it is not necessary, except in relation to secular variation, to attempt to attain the customary accuracy.

Distribution of Magnetic Stations. About 7000 stations have been occupied in the United States of which perhaps 4000 were well marked stations with all three elements observed, and possibly 1500 of these are now usable. Many of the stations were occupied during the period of expansion of the survey, 1899–1918, and have disappeared or are no longer magnetically suitable. To indicate present conditions, an examination within the past year indicates that there is an area of about 15,000 square miles in California and Nevada in which all previous repeat stations have become useless.

Compared with many countries of Europe prior to 1939 the survey is quite inadequate in certain regions. Chapman and Bartels (2) list a first-order survey as having stations 20–40 km. apart. For secular change, the past practice has been to occupy every 5 years from 150 to 200 stations scattered more or less uniformly in all parts of the country. These are necessarily occupied at various times during the five-year period. The ideal plan was that followed by Denmark of occupying about 8 stations at the close of a five-year period but this is not practicable with an area of 3,000,000 square miles, not to mention Alaska and other noncontinental areas which require determination of secular change.

There are variations in the length of the working season in different parts of so large an area and limited funds permit the use of not more than two field parties as a rule.

In Alaska the problems are similar, with the added difficulty that work must be done in a very short season, that magnetic changes are often large and rapid, that azimuths are more difficult to obtain, and finally transportation is difficult. During the past two years the ten-year repeat survey was made chiefly by aeroplane, including visit to Point Barrow where observations were made in June under truly Arctic conditions, as illustrated in Fig. 11. The scope of the survey is thus much extended but still is limited by the availability of landing fields. Lack of recent Polar expeditions has handicapped determination of secular change in northern Alaska.

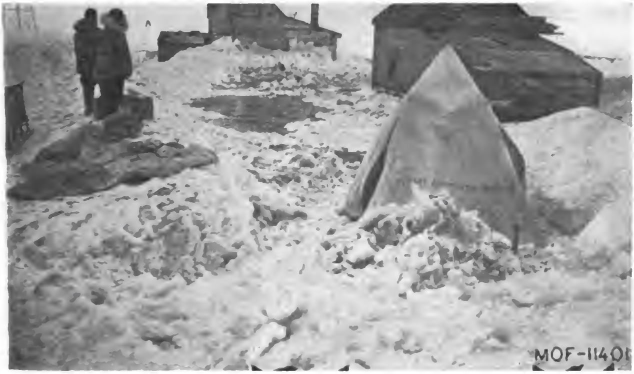


FIG. 11. Magnetic station at Point Barrow, Alaska.

In Alaska the practice of hydrographic and topographic parties (which operate in general along the coast) of making declination observations at frequent intervals is specially important and has led to the discovery and development of a number of large anomalies, both in area and intensity. Those that have been found indicate that there are many more in the great unexplored areas.

Puerto Rico and Hawaii have quite different weather conditions from Alaska and observations can be made at any time. However, in both cases, especially the latter, there are, because of volcanic formation, only a few places covered deep enough with coral to minimize the magnetic effects of lava. In both cases secular change is largely taken care of by the observatories which are located at undisturbed sites.

In Bache's time the magnetic survey at sea had not been conceived of except that vessels reported the magnetic variations during cruises—the results being more accurate than now obtained from the ordinary vessel because they were then built of wood. Values for the coastal charts were deduced as today from the land. It was recognized that serious errors were thus introduced, but the method remained unchanged till the cruises of the vessels of the Carnegie Institution of Washington—the partially non-magnetic *Galilee* and the non-magnetic *Carnegie*—gave really good values. An important step in the development was the work of Bauer and Faris aboard vessels of the Coast and Geodetic Survey.

When vessels of steel construction came into use, magnetic observations at sea were abandoned by the Coast and Geodetic Survey, except for investigation of areas of large anomaly.

During the cruises of the *Carnegie* enough observations were obtained in the United States waters to give fairly good control of the secular change; and various Arctic cruises, particularly those of Amundsen, performed the same function for Alaska. At present no observations at sea are being made or have been made since 1929, and the prospects of the *Research* becoming available are very slight. We have reverted to the earliest practice used in Bache's time of extrapolating curves based on shore observations. This is particularly unsatisfactory in regions of magnetic anomalies. For example, the *Carnegie's* observations in the general region of the Aleutian Islands did not fit in well with the shore results.

As a makeshift, the Coast and Geodetic Survey has made some declination observations in a whaleboat in the Gulf of Mexico with surveying vessel operating in the immediate vicinity. The results were fairly satisfactory, but application has to be limited to specially favorable conditions. The results are significant in connection with the surveys in the Baltic Sea by Sweden and Finland, which resulted in detailed surveys of all elements over a water area.

The data are made available in the form of maps and publications. Declination maps are issued for the United States and immediately adjacent parts of Canada and Mexico every five years, as for example, January 1, 1940, now in preparation. Maps of all the elements are issued every ten years. The repeat programs are carried out so as to provide the best possible data.

In spite of the impossibility of having the observations near the end of the period, an effort is made to approximate this, especially by making observations right through the course of the year for which the chart is to be published. This delays issue of the maps somewhat but is unavoidable if the chart is to have reasonably correct annual rates of change. Each observation is reduced to epoch as best possible and the lines of equal intensity (isomagnetic lines) and lines of equal change (isopors) can then be drawn. On the present maps, in spite of more observations, there is some smoothing so that the full complications are not shown. This can only be done on a very large scale map.

Uses of Geomagnetism. One of the earliest uses of geomagnetism on land is in connection with land surveys. Magnetic instruments were used to lay out Philadelphia; George Washington made land surveys with a magnetic compass; and the entire group of eastern and southeastern states was surveyed originally in this way. Laws of many states require county surveyors to test their compasses at regular intervals and such instruments are still much used. The Coast and Geodetic Survey, having compiled all available information as far back as 1750 for parts of the country, is constantly called upon to give the changes over specified periods. Even though misunderstanding of the complexities of secular change cause confusion, it is usually possible to find out how to rerun old lines and relocate missing corners. Compasses are much used for orientation and knowledge of approximate declination is useful. Declination and rates of change are given on all the aeronautical charts. The question has been raised as to how high anomalies have effect in distorting the average values and this is being investigated. At sea and on the Great Lakes compass data are needed for navigation. The normal values are given on all mariners' charts and disturbed areas are indicated. These tell the mariner where he cannot depend on his compass for navigation and other lesser anomalies where compass deviations should not be determined. Many charts and coast pilots give such information, especially in Alaska.

While the gyro compass is much used, it has not superseded the magnetic compass but rather increased its importance as the two devices are subject to different kinds of interference with their effectiveness.

The absolute vertical intensity data, while not having the same accuracy as the relative values, have several uses in connection with magnetic prospecting—first, to give an approximately correct absolute value, and second, to aid in arriving at the average value for a region on which to compute anomalies. There are important scientific uses of observatory data. Among these should be mentioned scientific knowledge of distribution of and changes in the earth's field for development of theories; and especially furnishing data regarding magnetic conditions useful in connection with radio transmission—especially the three-hour-range indices. Communication services, whether private or Government, make constant use of these data.

The maintenance of the magnetic survey requires constant struggle. Stations disappear or become disturbed; magnetic instruments deteriorate; magnetic material must be rigidly excluded from the vicinity of the instrument; and magnetic parts must be removed from the instruments themselves. Dampness must be kept to a reasonable minimum.

Among the important needs of the Survey are included: Replacement of equivalent of old survey; new stations to improve spacing; adoption of portable observatory plans and stations selected with special care; investigation of possible observation at night, when conditions are usually less disturbed, by use of light introduced without heat or magnetic effect; bringing and keeping publication of observatory results up to date; more research in regard to information given by observatories; a nonmagnetic survey vessel; a cooperative survey to provide data for new magnetic charts of the Caribbean (old chart of 1915 entirely out of date); precise absolute measurement of vertical intensity in observatory and field; and new instruments of modern design.

It seems that the attitude of Bache when the magnetic survey was new must still be maintained.

References

1. Vacquier, V. Short-time Magnetic Fluctuations of Local Character. *Terrestrial Magnetism and Electricity*, March 1937, pp. 17–28.
2. Chapman, S. and Bartels, J. *Geomagnetism*, Vol. I, Oxford, at the Clarendon Press, 1940.

THE SIGNIFICANCE OF FOSSIL MAGNETISM

A. G. McNISH

Department of Terrestrial Magnetism, Carnegie Institution of Washington

(Read February 14, 1941, in Symposium on Geomagnetism)

It has been said that less glory has accrued to the name of Alexander Dallas Bache because his work was connected with the Earth-sciences, these sciences being less fruitful and spectacular than other branches of research. This statement is only partly true as it leads to incorrect inferences regarding the importance of geophysics to the advancement of knowledge and the application of this knowledge to the betterment of mankind.

Truly, advances in geophysics are less spectacular than those in many departments of physics. They are the result of lengthy, painstaking efforts distributed faithfully over many decades. The geophysicist may not hope by means of a clever, brief experiment to discover a new and fundamental law of physics (which in a few years may be proved false by another clever, brief experiment). Yet there are problems in geophysics of such extent and profundity that solution of one of these would assure everlasting fame for the fortunate investigator who might accomplish it.

Geophysics has not been unfruitful in contributing to the advancement of mankind. As an example of a great contribution we have the magnetic compass which has guided the travels of man for many centuries on land and on sea, and now even in the air, and which remains today one of the most important of navigational instruments in spite of the strong competition being offered it by the far more elaborate and expensive gyrocompass. That the magnetic compass has faults one must admit, but these faults are far more frequently due to careless construction and installation than to defects inherent in its basic principle. The mariner's compass is a precise instrument for navigation; usually it receives less care and attention than a cheap alarm clock, and yet it performs its function faithfully.

In order that the magnetic compass may attain its maximum utility, a knowledge of the direction of the magnetic force at all places on the Earth's surface is necessary, and, since this force is slowly but constantly changing in direction and intensity, continual observations of it must be maintained. Although observations extend over four centuries, no definite law of these changes has been discovered—the changes are too complicated, and perspective too limited to admit even today of much more than a descriptive model of the phenomena. And this again emphasizes the long-time requirements of geomagnetic research. This meeting is being held to commemorate the initiation of systematic magnetic observations in this country by Alexander Dallas Bache one century ago in the heroic age of geomagnetism when such immortals as Gauss and Sabine pondered on the causes and found no answer. Today, with the additional data compiled during the last century, we still do not know. But one fact is evident—never in the present age of science will the complete picture of the gradual changes of geomagnetism unfold before our eyes.

Is our problem, then, hopeless of solution? We think not, for the geomagnetician has taken a hint from the biologist who, being unable to observe directly the evolution of the various forms of life, seeks the necessary evidence in the fossil remains of past ages. Just as rocks preserve a record of the evolution of various forms of life, so also they seem to have preserved a record of the evolution of the Earth's magnetic field.

This investigation is not new. Half a century ago studies were made on ancient bricks and potteries and on the magnetization of volcanic rocks. The results of the studies on volcanic rocks led to weird conclusions, the weirdest of which was that in a not too distant past the direction of the Earth's field was completely reversed. It is assumed that volcanic rocks, on cooling, become magnetized in a direction which coincides with the direction of the Earth's magnetism at that place when the rocks cooled. But in many cases the rocks may have become magnetic before they found their final resting places, thus leading to ambiguous results.

No suggestion that the direction of the Earth's field was completely reversed is to be found in historic records of geomagnetic observations. Reliable observations at a number of places extend back over four centuries. They show that large changes—as great as 35° in the direction of the horizontal com-

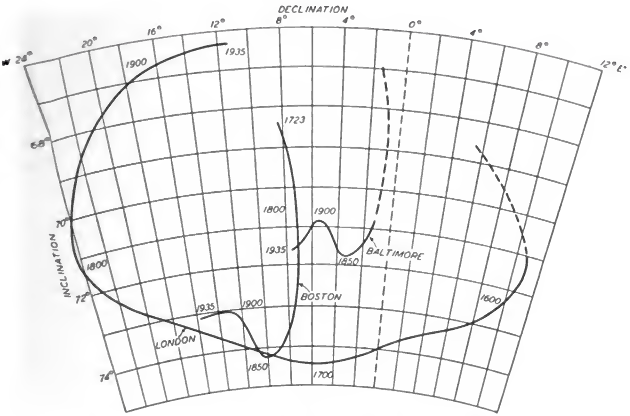


Fig. 1. **Secular variation in magnetic declination and inclination at London, Boston, and Baltimore.**

ponent in 200 years—have occurred. But there is no suggestion of a complete reversal in direction. A systematic study of the direction of the Earth's magnetism in past geologic ages based on the residual magnetism of rocks would clear up this important point and establish the necessary conditions which any theory of the geomagnetic field and its secular variation must fulfil. Fortunately suitable rocks are readily available for the prosecution of this investigation—the varved Pleistocene clays left at the retreat of the last glaciation. These clays are adaptable to this study for a number of reasons: They are unmetamorphosed and have retained essentially the same directions in geographic space ever since their deposition; they exhibit distinct layering, each layer corresponding to a year, so that the relative date of each layer may be determined; their chronology for 7,000 years has been reliably established by Antevs. The magnetic properties of these clays are due to the presence of finely-divided particles of ferromagnetic material, presumably magnetite which abounds in the hills over which the ice-sheets passed. During the summer thaws these particles were washed into the calm waters of the terminal lakes where they settled to the bottom. Large particles being deposited at first and finer particles during the ensuing

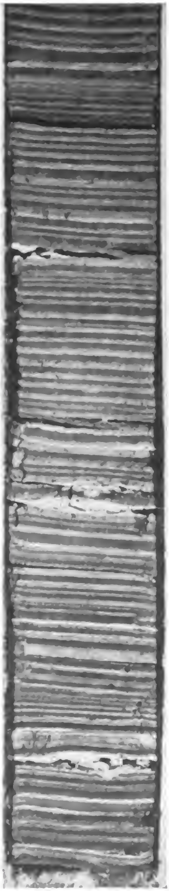


Fig. 2. Sample of varved clay.

winter gave rise to the distinct varving. The magnetic particles, as they settled to the bottom, aligned themselves with the Earth's field at that time, behaving like minute compasses. Subsequent depositions locked these particles in place, thus preserving a record of the direction of the magnetic field at the time.

These clays, like most sedimentary rocks, are only slightly magnetic as compared with volcanic rocks. However, their magnetization may be readily measured by a newly developed technique which is due in the larger part to my colleague, E. A. Johnson, who has been working with me on investigating the magnetization of these clays.

After these clays have been collected—care being taken to avoid all proximity to magnetic materials throughout the process—they are cut into cubes, the faces of which are aligned in a known geographic direction *in situ*. The cubes are then rotated in a coil to determine their magnetization without exposing them to any artificial magnetic field—a fault inherent in the method of measurement of most previous investigators.

Results of these measurements have been highly consistent. Directions of magnetization in adjacent annual layers agree with each other to within the limits of observational accuracy, considering the difficulties involved in squaring off and properly aligning the faces of the cubes. The directions of magnetization in widely separated layers, corresponding to different years, may differ considerably. Agreement between separate specimens from the same layers has been very good, even though the sites from which the specimens were obtained were several miles apart. Measurements, which reveal similar results, have been made on core-samples obtained from the ocean-bottom. However, with these core-samples it has not been possible as yet to refer the observed direction of magnetization to a fixed geographic direction nor to establish either a relative or absolute time-scale.

Interpreting these measurements as observations on the direction of the Earth's magnetism in past geologic ages, we may conclude that, during the period investigated, the direction of the field was not remarkably different from what it is today, and that changes were taking place similar in magnitude and rate to those which have been observed during historic times.

One may appropriately raise a question as to whether or not the magnetization of these varves as measured today closely corresponds to the magnetization they acquired at formation. In support of the belief that they do is the fact that the coercive force of the clays is very high, so that an extremely strong magnetic

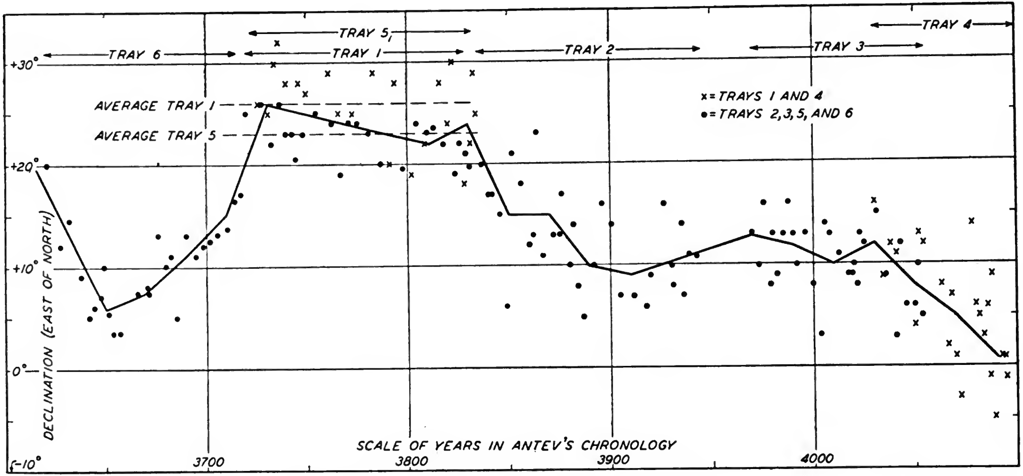


Fig. 3. Secular variation in declination, West Hartford, Connecticut, as deduced from the polarization of glacial varves; solid curve is drawn through 20-year averages.

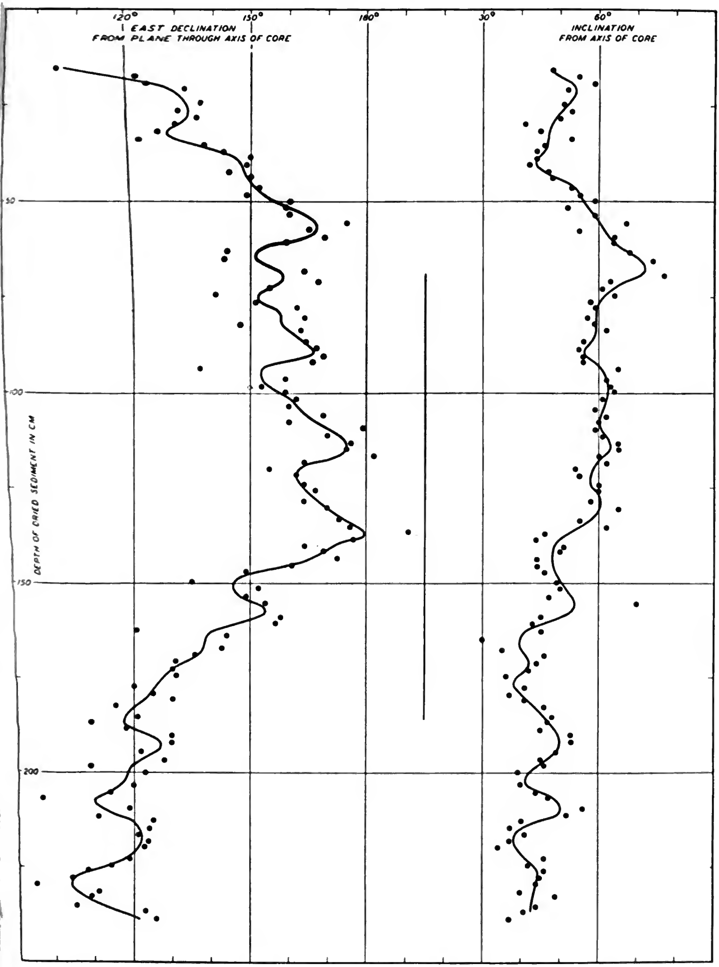


FIG. 4. Direction of magnetization of sediments in core No. 3 [curve smoothed by formula $(a + 2b + 3c + 2d + e)/9$].

field is required to alter their magnetization. The observations themselves supply the evidence that the clays are not randomly magnetized. If the Earth's magnetism had slowly altered their original magnetization during the thousands of years that they have lain in their present beds, tending to make the direction of their magnetization regress toward the present direction of the Earth's field, large differences in the direction of magnetization for layers differing in ages by one or two hundred years in ten thousand would not be expected. In short, by a process of exhaustion of possible causes, one is compelled to the belief that only the Earth's magnetic field at the time of their formation could be responsible for the directions of magnetization observed in the clays today, and hence that they portray to a high degree of fidelity the changes in the Earth's magnetism during past geologic ages.

Information derived from the study of fossil magnetism permits a more complete description of the Earth's field and its changes than has been possible before—a description which is essential to the formulation of a theory regarding its cause and origin. If, as was suggested by observations on the magnetization of volcanic rocks, the Earth's field has been completely reversed in past geologic times, one type of theory must be proposed. But if, as is supported by the evidence from the varved clays, the Earth's field has remained relatively constant throughout time, changing only in minor particulars, then a different type of theory is required. The evidence furnished by the varved clays fits well into a general concept derived from our present knowledge regarding the spacial distribution of the field and of its secular change.

Extensive data on secular change have been compiled by Fisk. He demonstrated that secular change is essentially a regional phenomenon. There are large areas, continental in extent, where the magnetic force is increasing and corresponding areas where it is decreasing. The most satisfactory representation of these changes on a world-wide scale is obtained by an isoporic chart for vertical intensity, that is, a chart on which appear contour-lines of equal annual change in that element. Examination of this chart reveals several centers where the rate of change is a maximum. Bearing in mind the distinctive features of this chart, attention may be directed to another.

As is well known, the Earth's magnetism may be represented to a fair approximation by the magnetic field of a uniformly magnetized sphere, or, what is its equivalent, the magnetic field of a small but powerful magnetic dipole at the center of this sphere. Such a representation accounts for only 80 per cent of the field. The remaining 20 per cent, or residual field, when plotted shows a striking relationship to the secular-variation field. This is best seen in vertical intensity. There is not an exact correspondence between the secular-variation field and residual field—maxima in the secular-variation field, that is, regions where the field is changing most rapidly, occur near the zero-lines of the residual field and zero-lines in the secular-variation field pass near maxima in the residual field. The dominant features of both fields are similar in extent and general form.

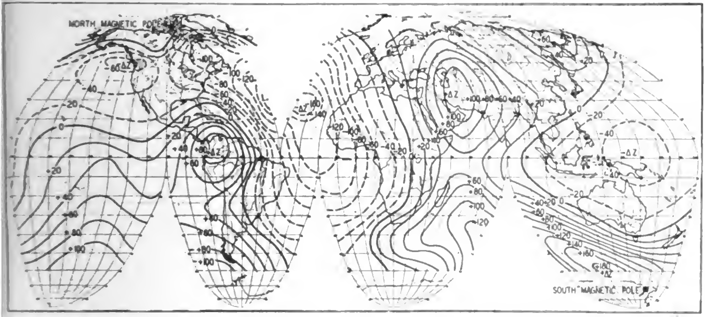


Fig. 5. Isoporic chart for vertical intensity (lines of equal annual change) approximate epoch 1920–25 (position of isopors in high latitudes, especially near the magnetic poles, very uncertain).

Both the residual field and the secular-variation field may be reproduced to within the accuracy of the observations by a few appropriately disposed dipoles. Although such a representation is not unique, a set of optimum positions can be determined so that the smallest possible number of dipoles is employed. Simplifying this model as much as possible, all these dipoles are located at the same depth, midway between the center and the surface of the Earth, and have their axes directed along radii. Furthermore, all the dipoles of the secular-variation field are of equal strength,

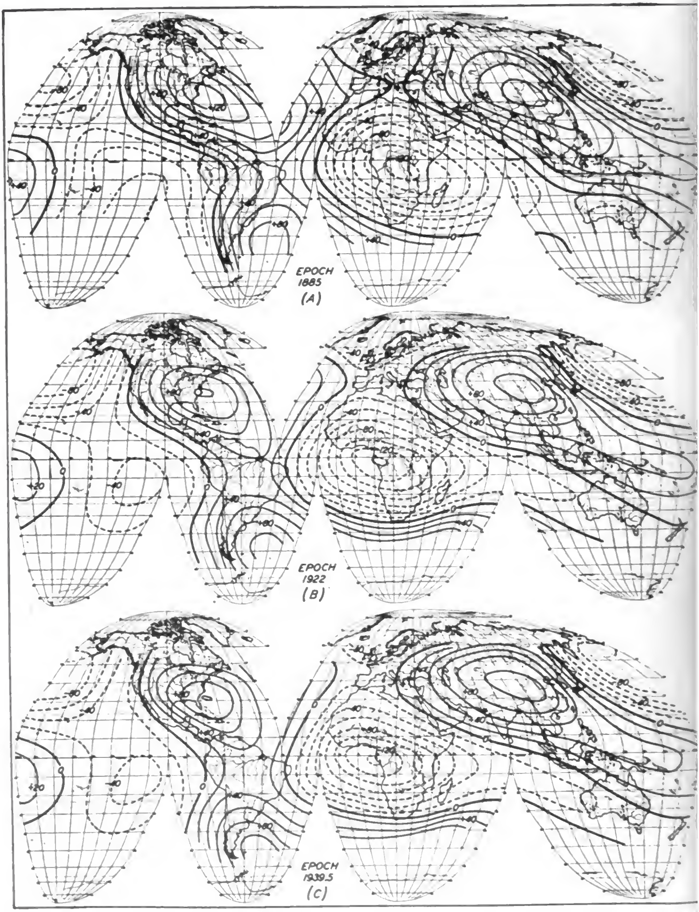


Fig. 6. Observed residual field in vertical magnetic intensity (observed field minus eccentric-dipole field); contour-interval 2000γ (2×10^{-2}CGS).

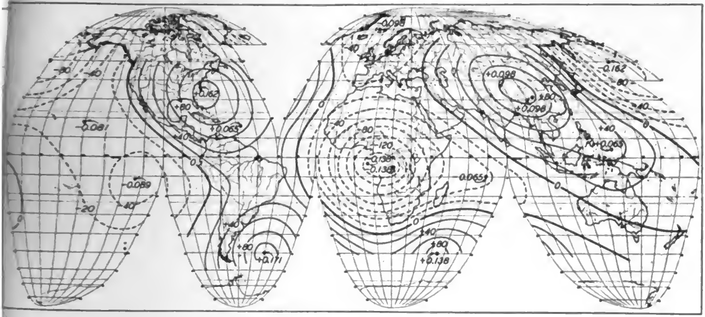


FIG. 7. Reproduction of residual field, epoch 1922, in vertical magnetic intensity by 14 radial dipoles on surface of Earth's core at locations marked by dots, strength of dipoles given in 10^{25} CGS [+ indicates north-seeking end downward; contour-interval 2000γ (2×10^{-2} CGS)].

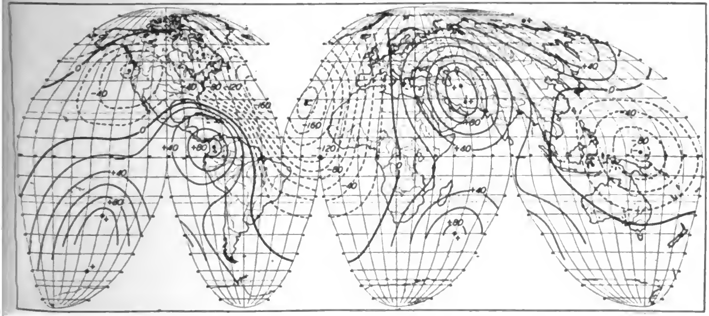


FIG. 8. Reproduction of yearly secular variation, epoch 1920–25, vertical magnetic intensity by 13 radial dipoles of uniform strength (0.137×10^{23} CGS) on surface of Earth's core at locations shown by dots [+ indicates north-seeking end downward; contour-interval 20γ (2×10^{-4} CGS)].

while the dipoles of the residual field are each about 100 times as strong as the secular-variation dipoles.

Stated in other words, the Earth's field and its secular variation at the present epoch may be represented by (1) a small but powerful magnet near the center of the Earth to produce the uniform field, (2) a set of 14 considerably weaker magnets appropriately located midway between the center and the surface (that is, at the discontinuity between the Earth's presumably fluid core and the mantle) with their north poles directed toward or away from the center, to produce the residual field, and (3) the yearly addition of a set of 13 magnets all of the same strength and about one-hundredth as strong as the magnets of the residual field and in general midway between them to produce the secular variation.

Clearly, the continual addition of the secular-change dipoles for a long time would build up a new residual field after about a century which would differ markedly from the one existing at the present time. Therefore, continuation of secular change for more than a century or two in the same sense seems unlikely—changes in secular change must occur. Historic records favor this view, since at no point do we have evidence that large secular changes have continued in the same sense for longer periods. Although the compass-direction at London changed 35°—from 11° east to 24° west—between the years 1600 and 1800, both before and after those years the change was in the opposite direction. Such is also the evidence supplied by the varved clays that, although large changes have occurred with relative rapidity, these changes have not continued in the same sense for a long enough time to alter the main aspects of the Earth's field.

This concept which has been presented is only a model. The Earth's magnetism cannot be due to concentrations of magnetic matter at individual points, but such concentration of matter at points constitutes a limiting case. Thus the residual-field dipoles at the surface of the Earth's core are symbolic representations of extensive regions of magnetization existing at lesser depths, and the secular-variation dipoles correspond to the growth or decay of magnetization on the edges of these systems. According to this concept, secular variation consists of a migration of the regions of magnetization which give rise to the residual field. Representation of these regions by dipoles at the surface of the Earth's core establishes their limiting lower depth—the regions themselves must be nearer the surface, namely, in the rocky mantle of the Earth. An excellent picture of the migration of those regions is revealed by magnetic charts of the residual field for various epochs. Distinct, progressive changes in location and extent of these regions are evident.

Thus, from the accumulated evidence of lengthy series of painstaking observations extending over several hundred years in time and over millions of square miles of the Earth's surface, a more coherent picture of the Earth's magnetic field and its secular variation is presented than was obtainable before. The general features of this picture are supported by the evidence derived from the study of magnetization in varved clays which permits extension of the picture backward into past geologic ages.. Further, and perhaps new and vital evidence will develop with the continuance of the investigation of fossil magnetism.

GEOMAGNETIC OBSERVATORIES AND INSTRUMENTS

H. E. McCOMB

Chief, Section of Observatories and Equipment, United States
Coast and Geodetic Survey

(Read February 14, 1941, in Symposium on Geomagnetism)

The first magnetic observatory in the United States was erected at Girard College, Philadelphia, in 1838 under the direction of Alexander Dallas Bache. Systematic magnetic observations in this observatory were started in May 1840 and were continued for six years. With his usual interest, care, and foresight in scientific work, Professor Bache exercised every reasonable precaution to see that the observatory and auxiliary equipment were constructed of nonmagnetic materials. Using eye-reading instruments, designed by Gauss and constructed in Göttingen, observations were made every hour throughout the day for magnetic declination and horizontal intensity, with some interruptions, of course. Later, a vertical intensity instrument, designed by Professor Lloyd, was obtained and placed in operation. The plan of the observatory and the arrangement of the instruments are shown in Fig. 1. Compared with modern observatory variometers, the instruments were quite massive. The magnet of the horizontal-component instrument weighed 25 pounds and was oriented and held in the magnetic prime vertical by torque applied through a bi-filar suspension composed of bundles of silk fibers. In this early work particular attention was given to the matter of proper orientation of the magnets responding to the different components.

During the term of office of Bache as Superintendent of the Coast Survey, the results of all the observations made at Girard College Observatory and the results of the early magnetic survey of Pennsylvania were published in the annual reports of the Superintendent for 1859 and 1862. In some of this field work there was used a large declinometer which has come to be known as the Bache magnetometer (Fig. 2). It is of interest to compare

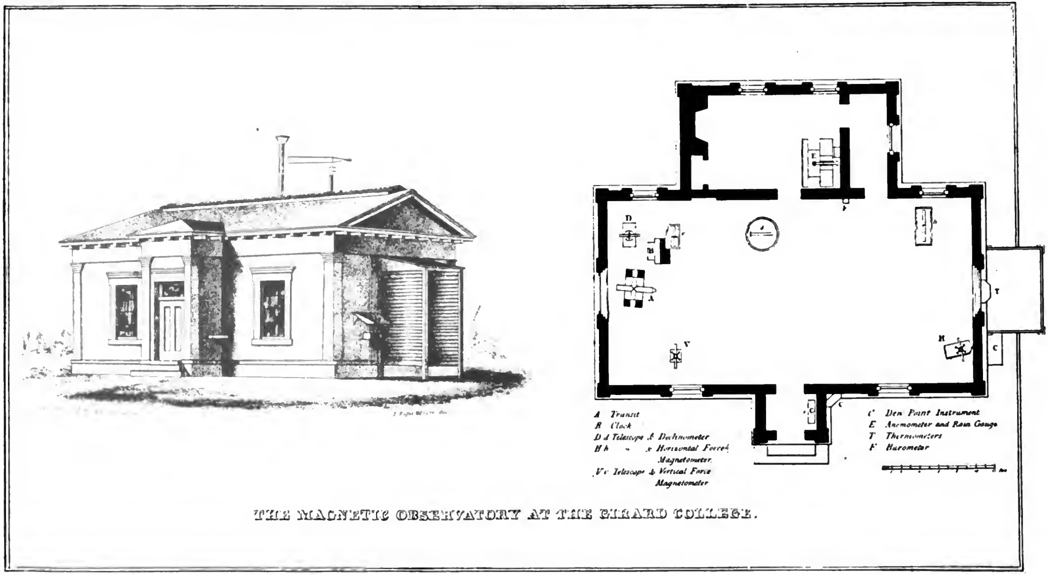

A Transit
B Clock
D d Telescope & Declinometer
H h " " & Horizontal Force Magnetometer.
V v Telescope & Vertical Force Magnetometer
C Dew Point Instrument
E Anemometer and Rain Gauge
T Thermometers
F Barometer
THE MAGNETIC OBSERVATORY AT THE GIRARD COLLEGE.

Fig. 1. Magnetic observatory at Girard College, Philadelphia, showing general arrangement of piers and instruments and view of the exterior. (Copy of plate in "*Magnetic and Meteorological Observations*, Girard College, Philadelphia, 1840–45.")

this instrument with the theodolite magnetometer designed and constructed by the Coast and Geodetic Survey in 1891 (Fig. 3).

In this early magnetic work it soon became evident that it would be quite desirable to provide self-recording instruments, and in this way eliminate much of the strenuous observing program and at the same time provide more complete pictures of the changes in the magnetic elements. Accordingly, an observatory for this purpose was constructed on the grounds of the Smithsonian Institution in Washington in 1852 and the first recording variometer, known then as a self-registering declinometer, was installed and placed in experimental operation the following year.

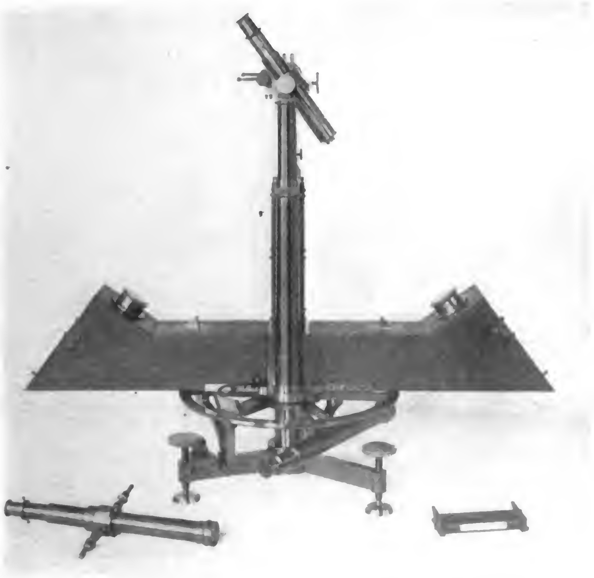


FIG. 2. Bache Magnetometer as used by Professor Bache from 1840 to 1843 in his magnetic survey of Pennsylvania. Instrument manufactured by Gambey of Paris. Photo by Dept. Terr. Mag., C. I. W.

These same instruments, originally constructed by Mr. Charles Brooke of London, were somewhat improved and installed in a magnetic observatory in Key West, Florida, where a rather complete series of photographic records was obtained from 1860 to 1866. The photographic paper was sensitized at the observatory as needed. An attempt was made to provide temperature compensation on the intensity instruments. For the bifilar horizontal intensity instrument, the compensating device, consisting of tubes of zinc and glass, was attached between the suspensions and the magnet in such a manner that the distance between the two bundles of fibers of the bifilar suspension varied continuously

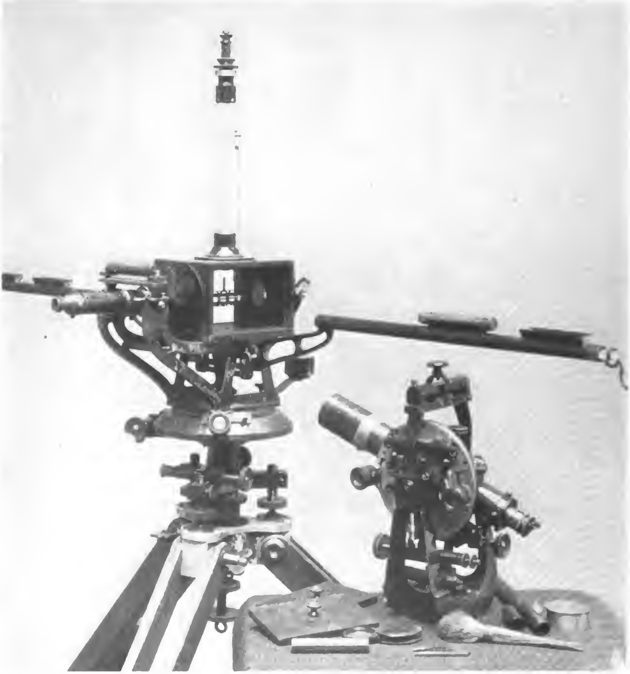


FIG. 3. Theodolite magnetometer designed and constructed by the U. S. Coast and Geodetic Survey in 1891 for field observations.

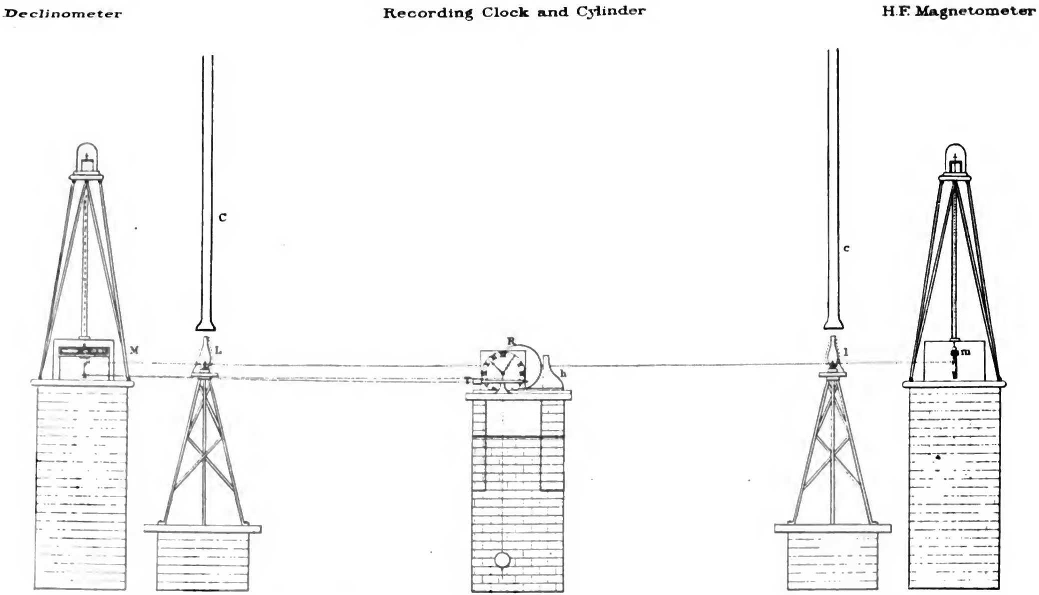


Fig. 4. Brooke magnetograph as operated first at the Smithsonian Institution in Washington in 1852 and after some modification installed and operated from 1860 to 1866 in the magnetic observatory at Key West, Florida.

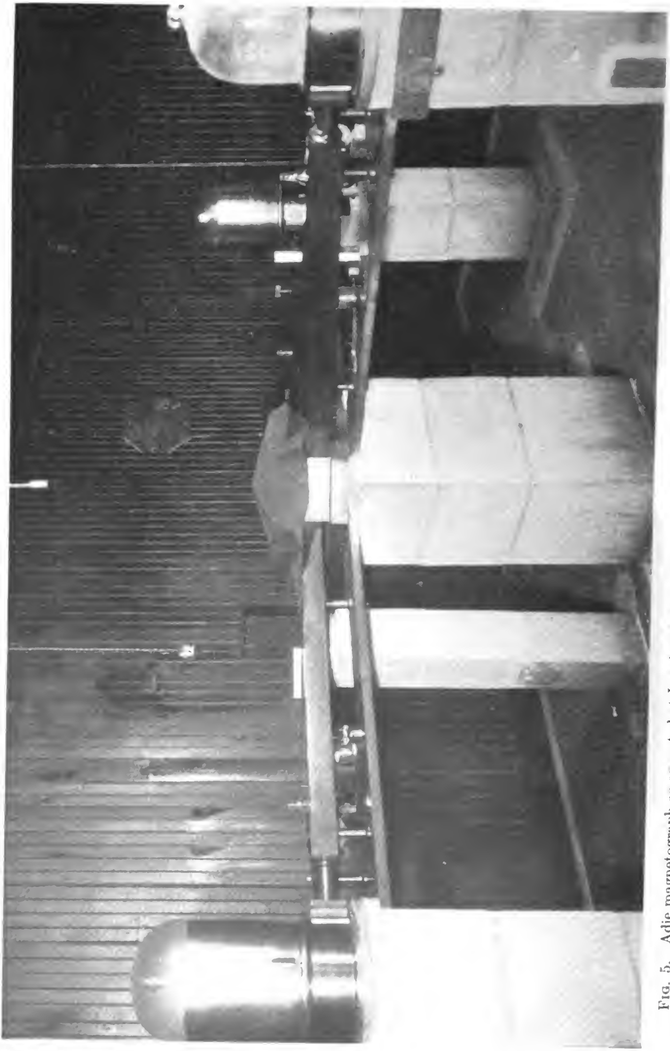


FIG. 5. Adie magnetograph as operated at Los Angeles, California, later at San Antonio, Texas, and now at Cheltenham, Maryland

as the temperature changed. This was contrived by utilizing the differential expansion of the glass and the zinc. The resulting changes in torque of the suspension were supposed to compensate for the changes in the magnetic moment of the suspended magnet, but the method did not prove to be satisfactory. A mercury thermometer attached to the recording magnet of the vertical-component variometer was supposed to provide temperature compensation for this element, but it too, was not of practical value.

In 1877 the magnetograph formerly in operation at Key West was installed at the University of Wisconsin at Madison, where an observatory was maintained until 1881.

For various periods between 1860 and 1883 observatories equipped with eye-reading instruments were operated intermittently at Eastport, Maine; Portland, Maine; Washington, D. C.; Point Barrow, Alaska; and at Fort Conger, Grinnell Land.

In 1882 a new magnetic observatory was constructed by the Survey on the grounds of the Branch State Normal School in Los Angeles, California, and the following year an Adie magnetograph, purchased during the Civil War (Fig. 5) was placed in operation

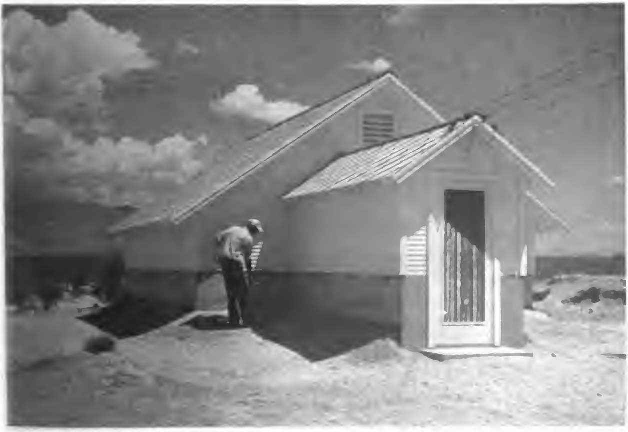


FIG. 6. Magnetic variation observatory at Tucson, Arizona.

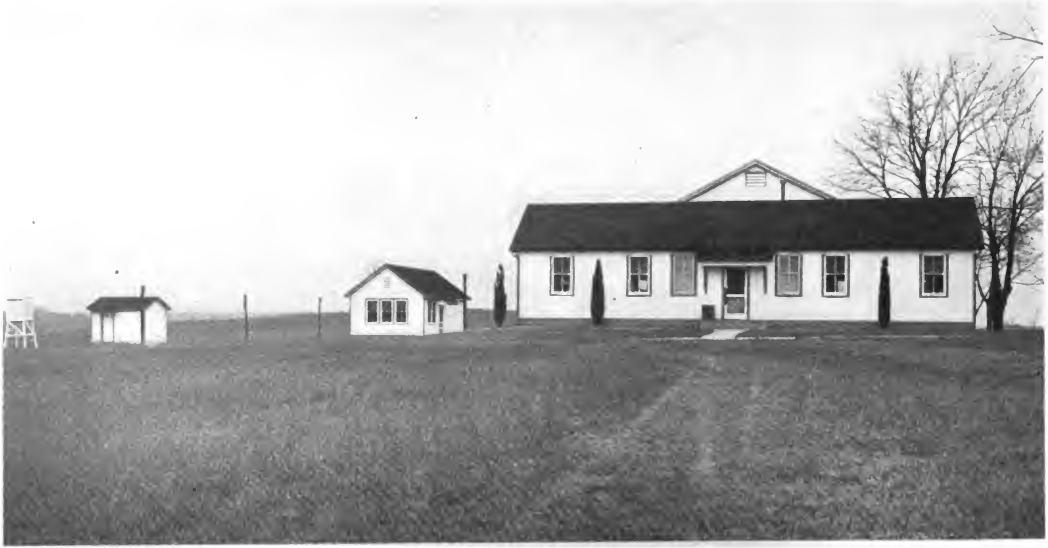


FIG. 7. Cheltenham magnetic observatory, Cheltenham, Md.

and continuous records obtained for about seven years. The instruments were then moved to a new observatory in San Antonio, Texas, where they were operated from 1890 to 1895 at two different sites in succession. The baneful effects of ground currents from an electric railway system were here well brought out.

Through the efforts of Bauer, who was in charge of the magnetic work of the Coast and Geodetic Survey from 1899 to 1904, five magnetic observatories equipped with improved magnetographs were established in various parts of the United States and its possessions. When he became Director of the Department of Terrestrial Magnetism of the Carnegie Institution of Washington in 1904, that Institution established observatories in Peru and Australia. In the meantime, of course, the magnetic work had been expanding in other nations all over the world so that today, under normal conditions, there are in operation about 60 permanent or sub-permanent magnetic observatories. In the construction of these modern observatories, considerable attention is given to the matter of insulation against temperature changes as

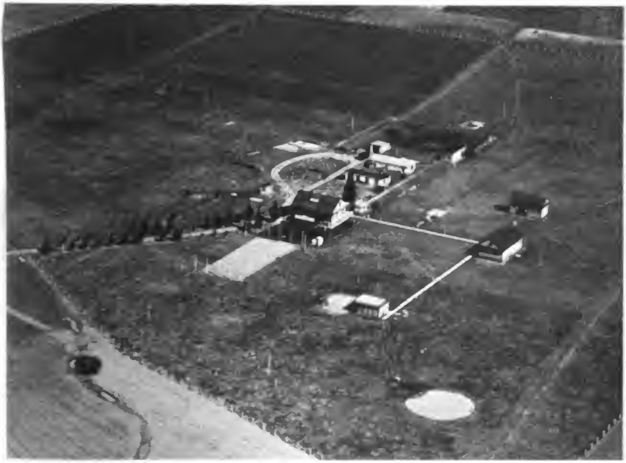


FIG. 8. Aerial view of Huancayo magnetic observatory, Huancayo, Peru. (Courtesy of Dept. Terr. Mag., C. I. W.)

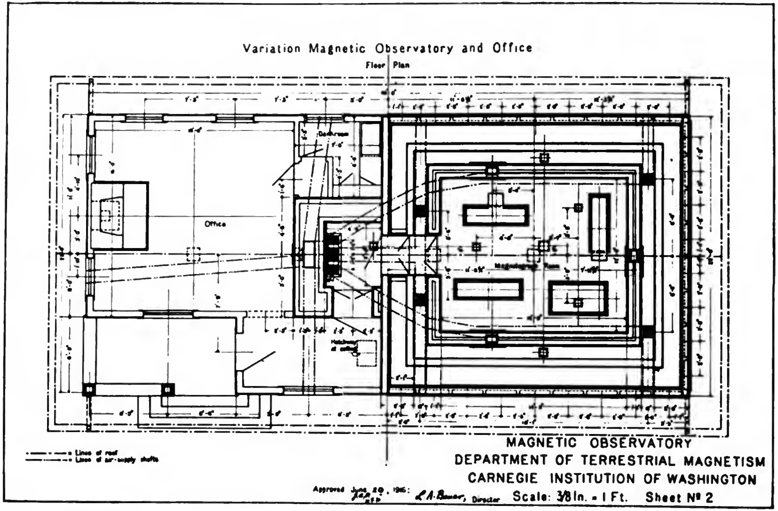

Variation Magnetic Observatory and Office
Floor Plan
Office
Darkroom
Magnetograph Room
Lines of roof
Lines of air-supply shafts
MAGNETIC OBSERVATORY
DEPARTMENT OF TERRESTRIAL MAGNETISM
CARNEGIE INSTITUTION OF WASHINGTON
Approved June 20, 1916: L. A. Bauer, Director
Scale: 3/8 In. = 1 Ft.
Sheet No 2

FIG. 9. Detailed plan of magnetic variation room for housing a complete magnetograph. (Courtesy of Dept. Terr. Mag., C. I. W.)

well as to provide effective temperature-compensating devices on the intensity variometers (Fig. 9).

Except for the field instruments designed and constructed by the Survey and by the Department of Terrestrial Magnetism of the Carnegie Institution of Washington for their observatories, practically all the observatory instruments have been obtained from Europe. Among these may be listed the Schulze earth inductor for reasonably accurate determination of magnetic dip (Fig. 10); the Cooke magnetometer (Fig. 11); the Eschenhagen magnetograph; the la Cour magnetograph; and the Askania field balance. One of the latest instruments for convenient and rapid comparison of horizontal intensity at two adjacent stations or at two stations remote from each other is the Quartz Horizontal Magnetometer, commonly called the QHM (Fig. 12). This instrument was designed by Dr. la Cour of the Danish Meteorological Institute. It is primarily a torsion instrument, the quartz fiber supporting the magnet being attached to the magnet holder and to the torsion head in a unique manner to prevent slipping of the fiber when under torsion (Fig. 13).

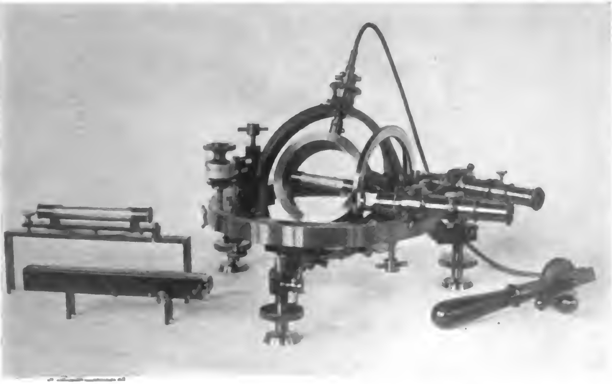


FIG. 10. Schulze Earth Inductor for use in measuring magnetic inclination. (Courtesy of Dept. Terr. Mag., C. I. W.)

For precise, routine measurements of horizontal intensity the

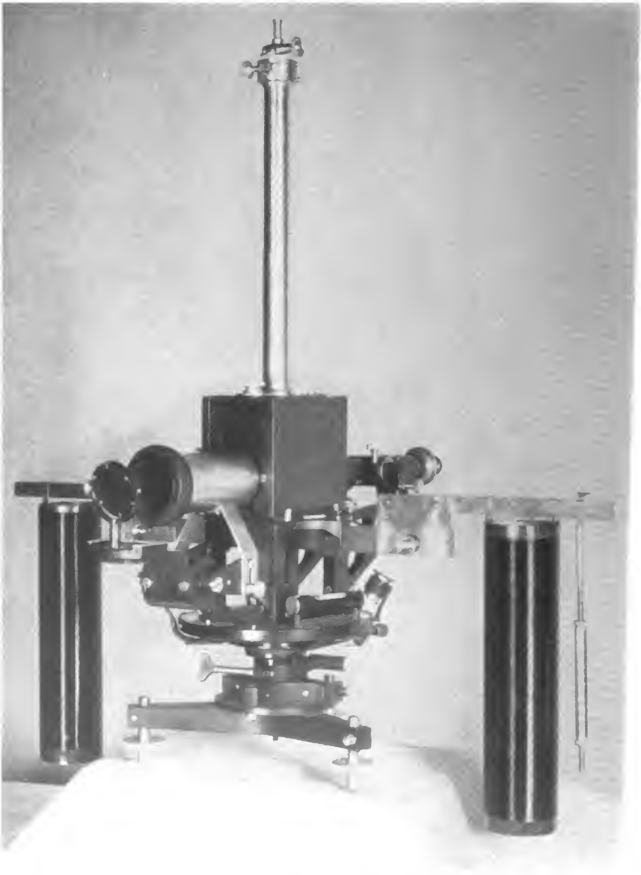


FIG. 11. Cooke magnetometer as modified by Coast and Geodetic Survey. Apparatus for determination of induction coefficients mounted on special deflection bar.

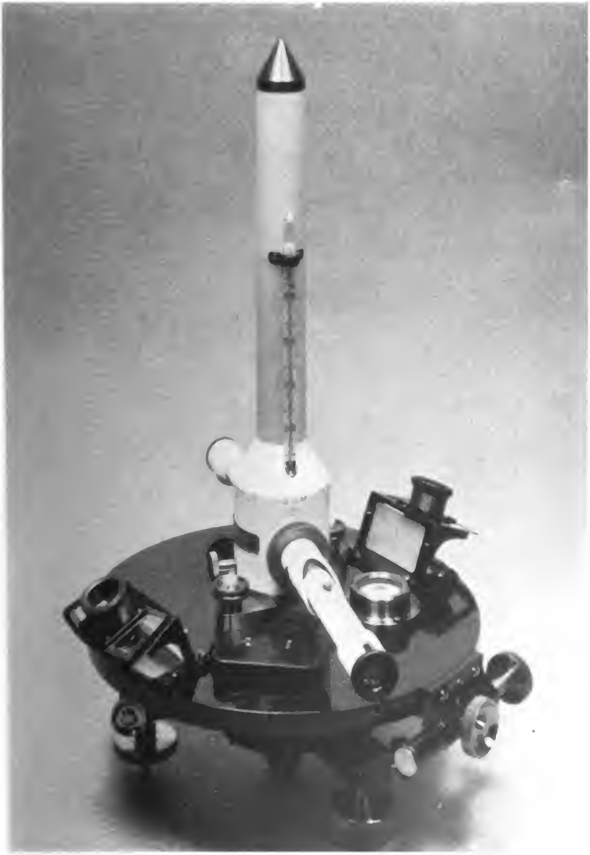


FIG. 12. Quartz horizontal magnetometer as designed by Dr. la Cour, Danish Meteorological Institute, and installed on special base by C. and G. Survey. Useful in determining relative values of horizontal intensity to a high degree of accuracy.

sine galvanometer, as designed and constructed several years ago by the Department of Terrestrial Magnetism of the Carnegie Institution of Washington, has been in routine operation at Cheltenham Observatory for about six years (Fig. 16). Similar instruments are in use at several foreign observatories.

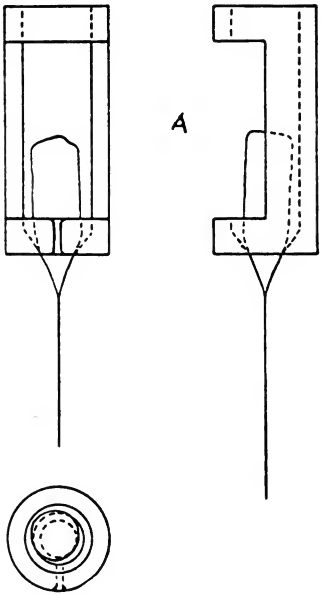


Fig. 13. Method of attachment of la Cour suspensions. The end of the quartz fiber is about one millimeter in diameter and the tapered portion fits into a conical cup. Friction is sufficient to prevent slipping of the fiber when placed under torsion.

It is possible to measure horizontal intensity with sufficient precision for all practical purposes. One of the remaining unsolved problems (at least in this country) is that of the absolute determination of vertical intensity to the same precision that is now possible for horizontal intensity. A precision instrument for this purpose is now under construction in the magnetic laboratory of

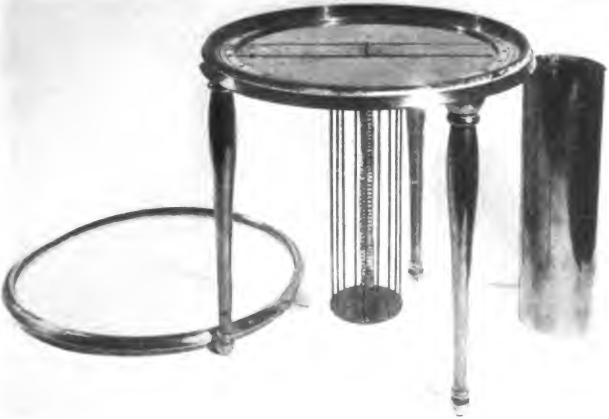


FIG. 14. Early type of relative intensity instrument, the torsion being supplied by a spiral spring mounted below the magnet. Torsion angles observed at lower end of the spiral spring. (Courtesy of the American Philosophical Society.)

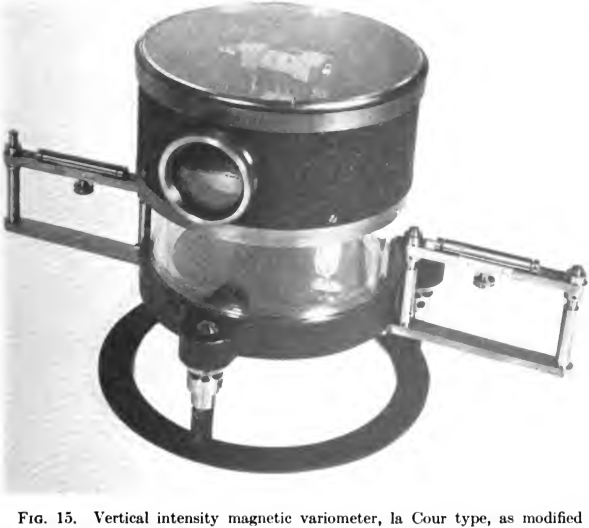


FIG. 15. Vertical intensity magnetic variometer, la Cour type, as modified by C. and G. Survey for magnetic temperature compensation and magnetic sensitivity control. Now in operation at Honolulu Magnetic Observatory.

the Department of Terrestrial Magnetism of the Carnegie Institution of Washington and it is hoped that it will be available soon.

At the Survey observatories, hourly values of H and Z are scaled directly in gammas and the values of D in minutes of arc,

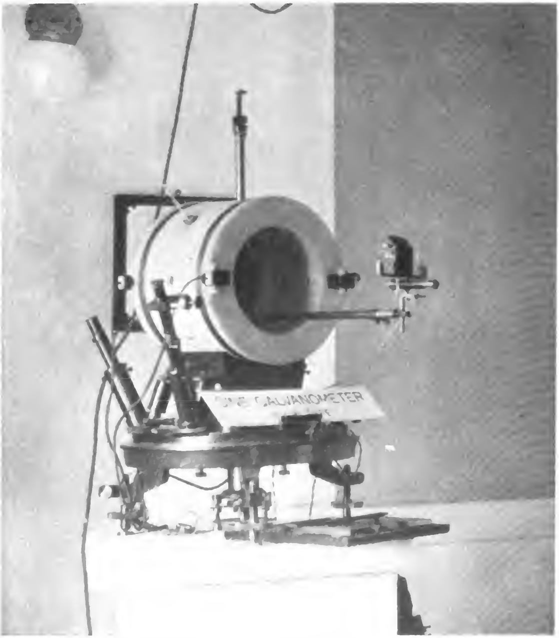


Fig. 16. Sine galvanometer for precise determination of horizontal intensity. Instrument constructed by Dept. Terr. Mag., C. I. W. and now in operation at Cheltenham Magnetic Observatory.

thus eliminating many steps in the reduction process (Fig. 17). Final results from magnetic observatories and results of observations made in the field are tabulated and published in form convenient for use by the public.

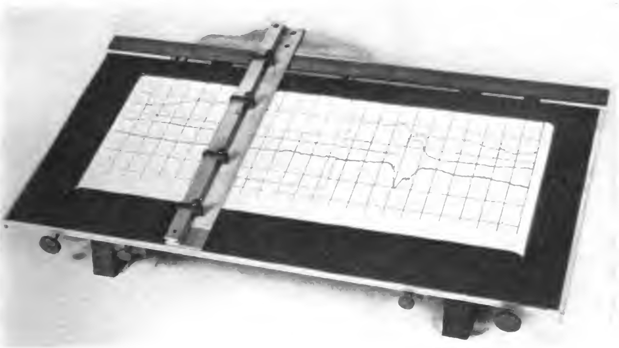


Fig. 17. Magnetogram mounted on special scaling board for direct scaling of magnetograms in gammas for the intensity records and minutes of arc for the declination records. In use at all Coast and Geodetic Survey magnetic observatories.

It would not be possible in a short paper to describe in detail all of the observatories and instruments which have been designed and used in the United States and in other countries. The few that have been described may serve, however, to illustrate that real progress is being made in geomagnetic measurements.

MAGNETIC WORK AT SEA

H. F. JOHNSTON

Chief, Section of Observatory Work, Department of Terrestrial Magnetism, Carnegie Institution of Washington

(Read February 14, 1941, in Symposium on Geomagnetism)

THE sea has always beckoned to adventurous men. Love of the sea inspired Homer to write that crowning nautical epic *The Odyssey,* wherein he tells of the fantastic voyages of his hero among the cave-dwelling giants, the Cyclops, of how he escaped from Circe and the fair sirens, and survived the mad grip of Scylla as she tried to engulf him in the swirling indraught of Charybdis.

But the Phenicians rather than the Greeks were the great sailors of ancient times. Their first recorded voyage is dated about the twenty-eighth century B.C. The inscriptions on the walls of Thebes give ample details of a maritime expedition in 1500 B.C., and, in the light of modern historical research, it seems fairly certain that they circumnavigated Africa about 600 B.C. The Romans sailed the Mediterranean in every direction in the time of Christ, fed their populace with grain from Africa, and carried their conquering legions in well-manned galleys, not only in the Mediterranean, but also across the English Channel.

The most extraordinary feats of navigation ever accomplished by adventurous voyagers will perhaps always remain conjectural. Whereas the Phenicians, Greeks, and Romans sailed on comparatively quiet inland seas or coasted along the shores of Arabia and Africa, the Polynesians of the same period spread eastward across the broad Pacific where distances are measured by thousands rather than by hundreds of miles. We must remember that these navigators made their way from place to place having only the sun and stars to direct them. They had to wait until about 1000 A.D. for a more precise instrument to give them direction. Though the magnetic properties of the loadstone had been known for at least 2500 years before the beginning of the Christian Era, it was not until the end of the eleventh century that Arab and Persian navigators used a crude form of compass, namely, a magnetized needle attached to a piece of wood and floated in a container on water. Proof of its use is found today in the extant treatise, corresponding to our modern encyclopedia prepared by the Chinese writer, Keou-Tsoung-Chy in 1111 A.D.

Compasses were in use in Western Europe by the year 1187 and the first compass with an attachment for determining azimuth was described by its inventor, Peter Peregrinus of Maricourt, in 1269. By 1403, compasses were so regularly used that an inventory of a vessel made at that time specifically mentions them. Soon after compasses came into regular use, the discovery was made that they did not indicate the true north and, as a result, some time about 1450 the makers of compasses attached the cards to the magnets in such a way that true instead of magnetic directions were indicated.

A sound foundation for magnetic work at sea was laid in the year 1457 when the first nautical almanac giving the declination of the sun for each day of the year was published. A few years later the Portuguese scientist, Zacuto, published his Almanac Perpetuum giving the sun's declination for the four-year cycle. Thus it was possible for Columbus on his voyage of discovery to the New World to measure magnetic declinations by his compass-observations of the direction of the sun at sunrise and sunset. He took "amplitudes" of the sun and our Practical Navigator Tables still contain a table of amplitude.

The first systematic magnetic survey at sea was completed by Halley in the years 1698–1700. From his observations of magnetic declination, he published the first chart in 1700 showing for the North and South Atlantic Oceans isomagnetic lines, that is, lines of equal declination. There is, however, a progressive change with the years in the direction in which a compass points at any one place on the Earth's surface as may be realized by enumerating the changes that have occurred at London, England, since the earliest recorded observations.

In 1580 the magnetic needle there pointed 11° east of north, its direction changed gradually to the westward until at about 1670 it pointed true north and by 1818, when it reached its greatest deviation, it was 24° west of north. Since that time, the magnetic needle has changed in an easterly direction until it is now only about 10° west of true north. From this example it may readily be seen that magnetic observations at sea are continually necessary so that accurate charts may be compiled for the use of navigators.

During Bache's time, accurate charts of declination could be prepared from the observations made on wooden sailing vessels. He lived in a stirring era for magnetic work, stimulated by the great advances made by Gauss in that branch of science. In the year 1835, Fox adapted the dip-circle which had been invented by Norman in 1576 for measurements of inclination at sea. Later, Lloyd in 1848, while Bache was Superintendent of the United States Coast Survey, devised a method for determining total intensity at sea using a dip-circle, and for the first time it was possible to measure all elements of the Earth's magnetic field.

The United States Coast Survey made observations for declination in coastal waters from early times, however, it was not until 1879–1880 that observations of intensity are recorded as having been made at sea off Cuba and Yucatan. By 1900 the survey vessel *Blake* was fitted with magnetic instruments for a sustained program of magnetic observations. The Department of Terrestrial Magnetism of the Carnegie Institution of Washington utilized the brigantine *Galilee* from 1905 to 1908 to obtain magnetic observations in the Pacific Ocean and in 1909 constructed the almost completely non-magnetic wooden vessel *Carnegie* for obsevations on the high seas. She was designed for her special purpose, there were glass-covered observing domes with removable panels for housing the instruments for measuring declination, intensity, and inclination. The design of all instrumental appliances was changed to insure the practicability of observations during the roughest seas and to eliminate errors due to improper orientation and lack of level.

Two magnetic instruments were mounted for measuring declination, one a liquid compass with azimuth-device for observing the sun's bearing directly, and the other a collimating-compass designed especially for measuring declination at sea. The accuracy of declination observations with an azimuth-device is not much better than half a degree, because of the continuous oscillatory motion of the compass-card in the horizontal plane and of the azimuth-attachment in the vertical plane. In the collimating-compass, the basic principle of which is due to William J. Peters, the direction-reference object, sun or star, is in view at all times, and hence the oscillatory motion of the compass-card is eliminated. Observations of declination with the marine collimator have random errors of only one- or two-tenths of a degree under the motion from moderate seas and three- to four-tenths of a degree under extreme motions.

In constructing the marine collimator a liquid compass was used, the card being removed and replaced by four 10°-scales at the north, south, east, and west points. Four concave mirrors of speculum were mounted on the float so that the scales were in their focal planes. The bowl was provided with four windows, thus permitting viewing of the scales when the bowl was correctly oriented. The windows in the bowl were segments of a spherical shell whose center was at the point of support of the optical systems. Hence the rocking of the bowl or the rotation produced by yawing did not alter the optical conditions. The bowl was suspended in a gimbal-system which had four circular openings. A cross-section of the bowl is shown in Fig. 1 in which is readily

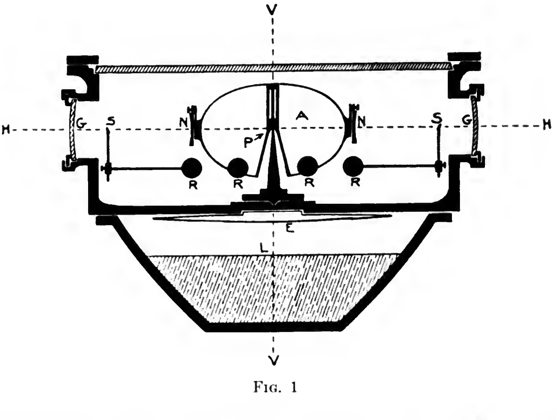


FIG. 1

discernible the position of scales, *S*, speculum mirrors, *N*, spherical windows, *G*, and magnets, *R*. Figure 2 shows the compass-float on which the speculum mirrors are mounted and the four 10°-degree scales. Figure 3 shows the bowl mounted in the gimbal-ring which has four circular openings for viewing the scales. In practice the angle between the central division of the scale and the celestial object is measured by means of a sextant. The image of the celestial object is brought on the scale and it is observed for several oscillations so that the amplitude may be determined.

Horizontal intensity was measured by a simple deflection-apparatus in which the magnet is mounted vertically above a compass-card and the angle of deflection is determined by comparison with an undeflected compass. If the value of the deflection-angle is u, then the value of the horizontal intensity, H, is given by the formula $H = (mC/\sin u)$, where m is the magnetic moment of the deflecting magnet and C is a function of the deflection-distance and its changes with temperature and of the induction- and distribution-coefficients. The value of C is determined from observations on land. With this instrument, horizontal intensity may be measured with an accuracy of approximately one part in 1,000.

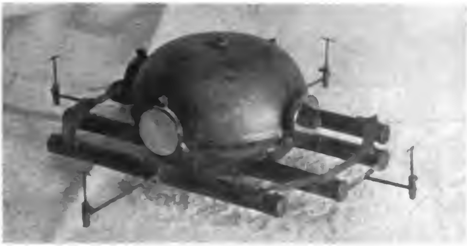


FIG. 2

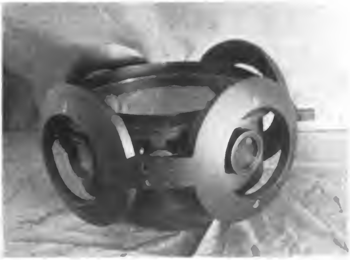


FIG. 3

Inclination was measured by two methods, in the first a dip-circle with semicircular agate cups for mounting the dip-needles and retaining them in the proper position during motions of the ship, and in the second an earth-inductor adapted for use at sea by an ingenious device which permitted rotation of the coil without disturbance to the level of the instrument. The induced currents produced in the coil when it was rotated in the earth's field were measured by a marine galvanometer of suitable sensitivity. In addition, an appropriately designed gimbal was made to mount it, in which the knife-edges were a hard alloy of platinum and iridium resting on slightly curved bearing pieces of agate. The outer gimbal-ring was made in two parts which could be rotated relatively to each other, thus making for ease in maintaining the earth-inductor in the magnetic meridian and also permitting 180°-reversal and tending to eliminate errors in level.

At times the utility of the magnetic compass has been threatened by improvements or modifications in the ships which carried them. Flinders and Airy invented correctors for compensating compasses on account of the iron which was used in greater and greater amounts in ships. Submarines of the largest size now use 8,000 amperes direct current when submerged with resultant large deviations in the magnetic compass and it was necessary to devise suitable electromagnetic correctors to compensate the error. The latest menace to magnetic compasses are the devices to combat magnetic mines, such a device as was mounted on the passenger steamer *Queen Elizabeth*, when she entered New York Harbor about a year ago.

Magnetic work at sea is still a necessity, so that up-to-date charts may be supplied to navigators. Since the loss of the *Carnegie* in 1929 very few observations have been made at sea. It is hoped that some time soon the non-magnetic vessel *Research*, of the British Admiralty, may begin her work at sea and so fulfill our magnetic needs.

GEOMAGNETISM: WORLD-WIDE AND COSMIC ASPECTS WITH ESPECIAL REFERENCE TO EARLY RESEARCH IN AMERICA

J. A. FLEMING

Director, Department of Terrestrial Magnetism, Carnegie Institution of Washington

(*Evening Lecture, February 14, 1941, in Symposium on Geomagnetism*)

THE complex nature of geomagnetism—the general magnetic field of the Earth—and of its varied phenomena has long puzzled and intrigued man. Despite several centuries of speculation and research, there is as yet no adequate explanation of how the Earth became magnetic or why it remains so. Closely associated with this problem and making it more difficult of solution is the "perpetual variability" of geomagnetism at any one place—changes, on the one hand, slow when compared with the life-span of the individual, as in its long-time or secular change, and, on the other hand, rapid, as in its short-time or more ephemeral changes; the latter are outstanding in their relations to cosmical phenomena in space about us. Thus the magnetic behavior of our planet can only be deduced from "the many and varied manifestations of terrestrial force in so far as it presents measurable relations in time and space."

To the layman seeing and hearing daily of navigation, the compass appears simple though mysterious in its directive ability which demonstrates that magnetic forces are present everywhere in the Earth. But to the scientific scholar that which is apparently simple is often the most baffling as in this case, perhaps not in the physical principles concerned, yet certainly in the origin and observed periodic and aperiodic fluctuations of the forces involved. Thus the distribution and variations of geomagnetism are intimately related not only to the structure of the Earth's interior and of its atmosphere but also to solar and cosmic phenomena.

We may well pay homage to the splendid and long-continued observation and research of our predecessors and of present-day scholars in all corners of the world which provide a rich and constantly growing mass of data. As compared with the more intensive observational work of the earlier years, experimental

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 2, MAY, 1941

approach, through investigations both outside and inside the Earth, is continually enlarging. There are involved not only the phenomena of the Earth's magnetic field, but their relations to other geophysical and astrophysical sciences.

HISTORY

Unlike many other natural phenomena, geomagnetism presents no ready evidence of its presence to any of the human senses. Thus its early perception through the directive property of a lodestone or magnet freely suspended in the Earth's field is veiled by the myths and legends of south-pointing chariots in China some 4,500 years ago and of its applications by the Egyptians, the Phenicians, the Greeks, and the Latins. We have, however, by the end of the twelfth century A.D., definite and well-authenticated evidence of the use of this property in navigation. Thence onward there was gradual transition from the field of speculation to that of scientific investigation and in 1600 Gilbert published his famous book on the magnet—the first treatise picturing the Earth as a great magnet. A century later Halley's world charts showing variations of the compass appeared. Wilcke's chart of magnetic dip or inclination was published in 1768. Charts delineating magnetic directive force resulted from Humboldt's observations on his American journeys during 1799 to 1803.

The period including the end of the eighteenth century and the first half of the nineteenth century was an era of unequaled constructive work in geomagnetism by many eminent scholars. Among these may be mentioned Humboldt, Gauss, and Lamont of Germany, Sabine and Airy of Great Britain, Poisson and Duperrey of France, Quetelet of Belgium, Hansteen of Norway, Kupffer of Russia, and Nicollet, Locke, Loomis, and Bache of America.

That thoughts on this subject were then not limited to scientific men is evidenced by a discourse of John Quincy Adams in our House of Representatives during preliminary steps bearing on the establishment of the Smithsonian Institution in which he said:

> What an unknown world of mind is yet teeming in the womb of time, to be revealed in tracing the causes of the sympathy between the magnet and the pole—that unseen, immaterial spirit, which walks with us through the most entangled forests, over the most interminable wilderness, and across every region of the pathless deep, by day, by night, in the calm serene of a cloudless sky, and in the howling of the hurricane or the typhoon. Who can witness the movements of that tremulous

needle, poised upon its center, still tending to the polar star, without feeling a thrill of amazement approaching to superstition?

It may be recalled that the man whose memory we honor in this centenary celebration was named in 1846 one of the regents of

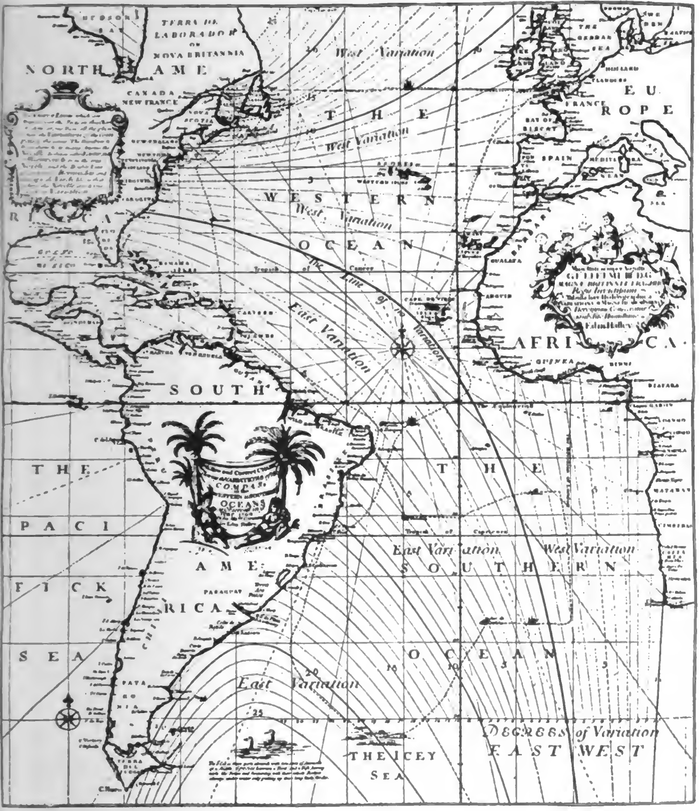
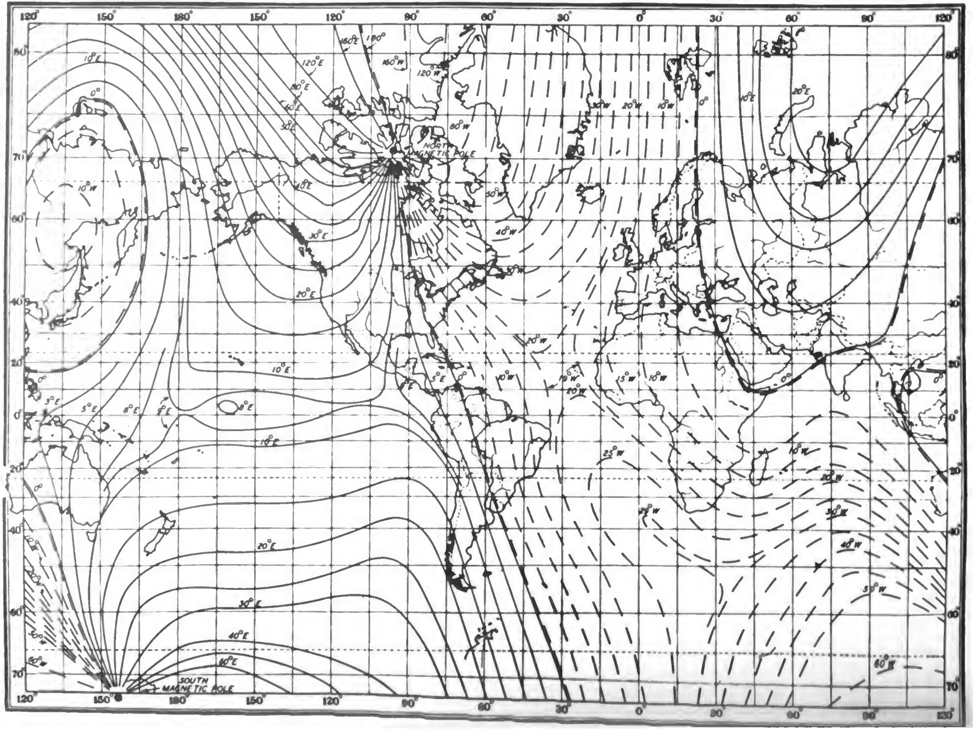


FIG. 1. Earliest isomagnetic chart of "magnetic variation" (that is, magnetic declination) by Halley for the Atlantic Ocean based on data during the cruise of his pink *Paramour* during 1697 to 1701.


120°
150°
180°
150°
120°
90°
60°
30°
0°
30°
60°
90°
120°
80°
70°
60°
40°
20°
0°
20°
40°
60°
70°
NORTH MAGNETIC POLE
SOUTH MAGNETIC POLE
10°E
0°
10°W
160°E
180°
160°W
120°W
120°E
80°E
60°E
50°E
80°W
30°W
20°W
10°W
10°E
20°E
40°E
50°W
40°W
30°E
20°E
10°E
5°E
0°
3°E
5°E
8°E
10°W
20°W
15°W
25°W
30°W
40°W
50°W
60°W
20°E
30°E
40°E
60°E

the Smithsonian Institution in the act of incorporation and that he was continued by the Congress in this office until his death—a period of nearly 20 years. It may be fairly said that his influence contributed much to bringing about the policy which has given that Institution its present high standing among scientific organizations of the world.

During this era it became evident that only through mutual assistance and international cooperation in physics, in geology, and in mathematics could it be hoped to advance understanding of geomagnetism. In this period were established the Magnetic Union of Göttingen for simultaneous observations all over the Globe and the British colonial observatories from which resulted the stimulation for greater activity in America by Bache and his colleagues. Through correspondence and a visit with Joseph Henry to Europe in 1837–38 Bache obtained instruments for, and took active part in, establishing the Magnetic Observatory at Girard College with help from its Trustees and the American Philosophical Society. To his enthusiam and that of his contemporaries and their pioneer work we are in great measure indebted for the prominent position now held by American organizations in geomagnetic research.

Bache became interested in geomagnetism soon after he accepted the Chair of Natural Philosophy and Chemistry at the University of Pennsylvania in Philadelphia (1828). He was not satisfied with the occupation of teacher alone and sought to extend the bounds of science by discoveries of his own. He became a member of the Franklin Institute and of the American Philosophical Society. He erected in his own dwelling an observatory in which, with the aid of his wife and of his assistant John Fraser, he determined for the first time in this country the periods of daily variations of the magnetic needle and, by a series of observations, the connection of the "fitful variations of the direction of the magnetic force with the appearance of the aurora borealis." In 1836, at the age of thirty, he was selected by the Trustees as President of the newly endowed Girard College for Orphans preparatory to organizing that institution. In this connection he was commissioned to go abroad to study systems of education, instruction, and discipline in Europe. You have already learned of his activities after his return in the interest of public education of Philadelphia and of his reinstatement in 1842 as Professor at the University of Pennsylvania. His visit in Europe had further increased his interest in geomagnetism and stimulated his cooperation in the enterprise of the British Association for the Advancement of Science and the British Government of establishing contemporaneous observatories at widely separated points to determine the magnetic elements. From this resulted the construction, through the Building Committee of Councils of the City of Philadelphia, of the Magnetic Observatory at Girard College to which the Trustees contributed instruments with latest improvements and which was supported by the American Philosophical Society as well as by many generous individuals. The observations there were continued for five years from May 1840 to June 1845. Expenses were first provided by the Girard College, then by the American Philosophical Society during 1841 to December 1842, and after a short interruption of the program during January to April 1843, by the Topographical Branch of the War Department during 1843 to 1845. Bache discussed the data obtained and published a long series of volumes thereon. His published contributions to geomagnetism began in 1832 with the appearance in the *Transactions of the American Philosophical Society* of his article "On the diurnal variation of the magnetic needle." Between 30 and 40 volumes were prepared by him during 1833 to 1865.

Bache's extensive geomagnetic discussions included in twelve large volumes of a memorable series called "Discussion of the magnetic and meteorological observations made at the Girard College Observatory, Philadelphia, in 1840, 1841, 1842, 1843, 1844, and 1845" is even today an epitome of avenues of research. Thus, for example, he has discussed the 11-year period in the amplitudes of solar diurnal variations and disturbances of magnetic declination, magnetic horizontal force, magnetic vertical force, magnetic total force, and magnetic dip; the annual inequalities of the elements; the influences of the Moon upon the several elements; and the effect of auroral lights on geomagnetic phenomena. Other papers by Bache relate to instrumental improvements, reports of surveys by himself and colleagues, comparison of results at Philadelphia, Toronto, and Hobarton, discussions of his magnetic observations on land and at stations in Europe and those made on the line of the boundary survey between the United States and Mexico during 1849 to 1852, and elsewhere in the United States.

It is not surprising that the Americas lagged behind Europe

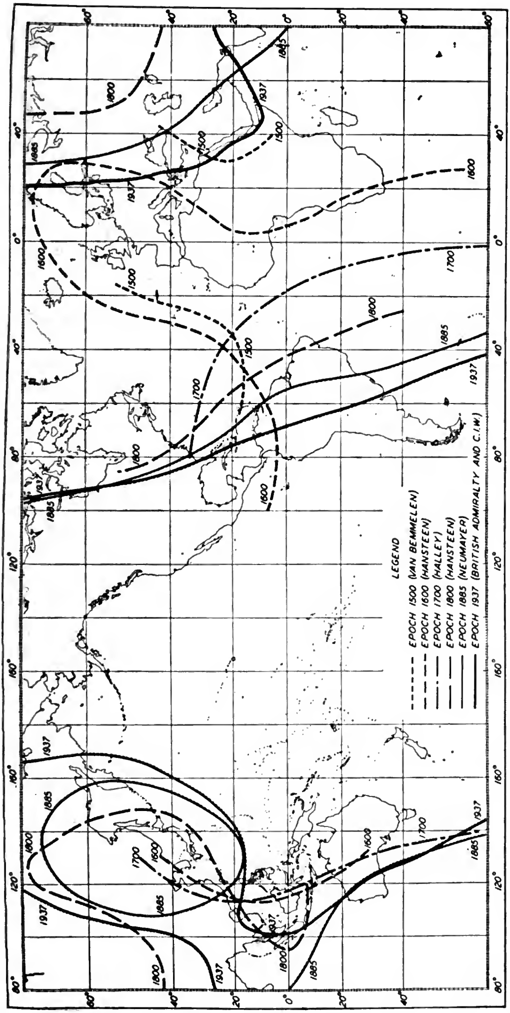


FIG. 3. Progressive change or secular variation of the lines of zero magnetic declination (agonics). [The dots indicate the side of the agonic lines on which the declinations were easterly.]

in geomagnetic investigations which had their inception there in earlier times—a fact still evidenced by the numerous magnetic observatories in the comparatively small area of Europe. Nevertheless, research in the western hemisphere figures prominently in the history of the Earth's sciences. One of the agonics—the line along which the compass points true north—passes through Uruguay, Brazil, Venezuela, Cuba, the United States of America, and Canada. The magnetic equator—the line joining points where the magnetic dip is zero—crosses Peru and Brazil. The delineation of the magnetic equator early aroused scientific interest

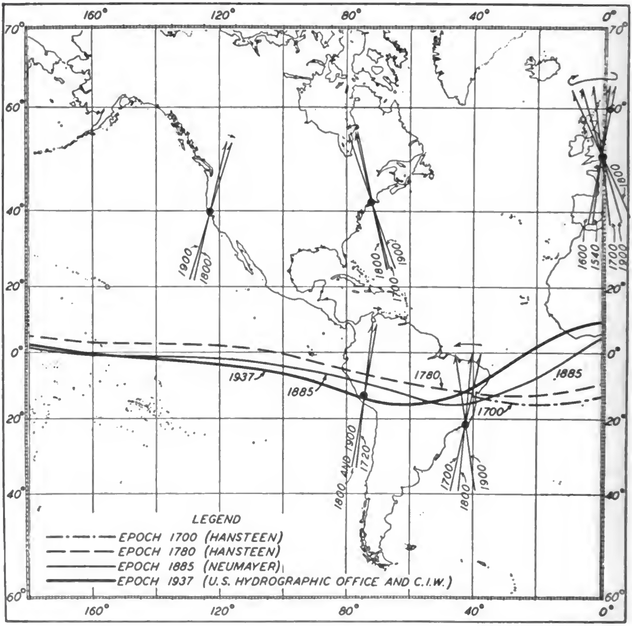


FIG. 4. Secular variation of the magnetic equator in the western hemisphere from 1700 to 1937 and secular variation of the magnetic declination or the compass-direction at London, Eastport, San Francisco, Callao, and Rio de Janeiro from earliest observed values.

since it deviates so greatly in South America to the south of the geographic equator. In 1802 Humboldt located it in the Andes near 7° south latitude between Cajamarca and the silver mines of Micuipampa—possibly the locality now known as Hualgoyoc. He

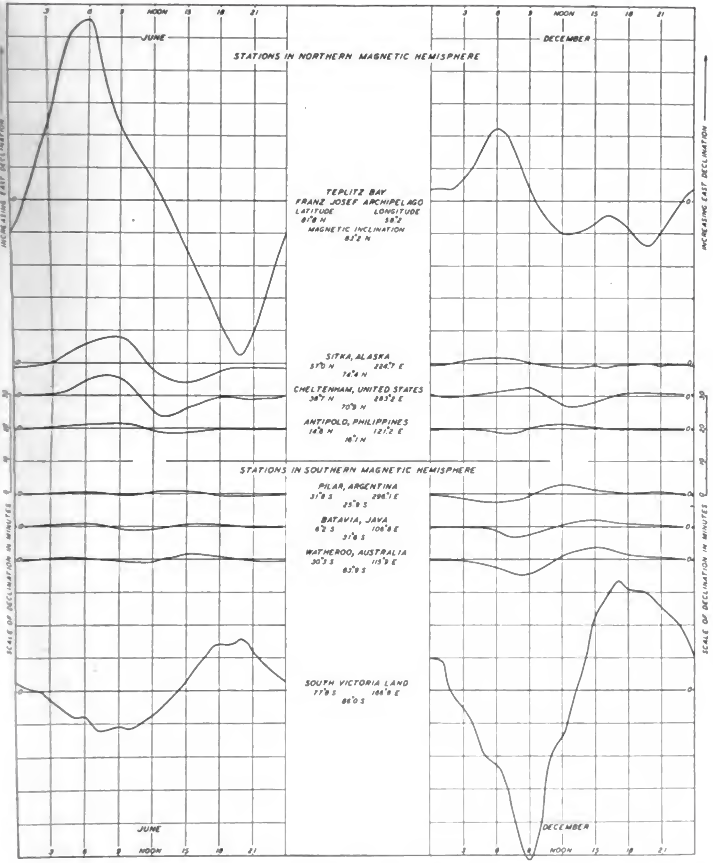


FIG. 5. Solar diurnal variation or inequality in magnetic declination [curves show seasonal changes and modification in ranges for different latitudes].

also measured the relative magnetic intensity there and, since he conceived that this intensity was lowest on the equator and increased toward the pole, he defined his observed value as unity on a scale of geomagnetic force. This scale was adopted by other observers, and measurements all over the Globe were expressed in it for several decades. With the later development by Gauss of means for absolute measures, this value was converted into the absolute unit, being found equal to 0.3494 centimeter-gram-second unit.

The earliest detailed State magnetic survey in North America was made by Bache in 1840 to 1843, namely, his "Magnetic survey of Pennsylvania and parts of adjacent States." The observations at 22 points were made by Bache during his summer vacations and at private expense. His isomagnetic map of that survey shows remarkable agreement, when account is taken of secular change, with the somewhat earlier map by Loomis of the eastern part of the United States based on observations made by Loomis in 1838 to 1841, by Nicollet in 1832 to 1836, and by Locke in 1838 to 1845. Nicollet occupied some 25 stations all in the eastern part of the United States. Locke obtained a series of magnetic observations at 100 stations from the State of Maine to some distance beyond the Mississippi and at the Magnetical Observatory of the British Government in Toronto at personal expense—a notable example of private research stimulated by the impulse and systematic direction which during this era had been given to this study of the Earth's magnetism. Locke's observations gave the first indication of the location of the North American focus of greatest magnetic intensity, which was verified by later investigations of Lefroy.

Bache's idea of intensive State magnetic surveys, in addition to the general magnetic survey of the whole country, was not realized until more recent years. For example, Professor Francis E. Nipher, of the Washington University at St. Louis, undertook a detailed magnetic survey of Missouri from 1878 to 1883; he also was dependent entirely upon private aid for defraying the expenses of this survey which included 149 stations. At about the same time some preliminary observations appear to have been made by Professor Gustav Hinricks in Iowa. These were followed in 1887 to 1890 by G. H. Cook in the survey of New Jersey, and in 1896 to 1900 by L. A. Bauer in the survey of Maryland under the direction of Professor W. B. Clark of Johns Hopkins University, State

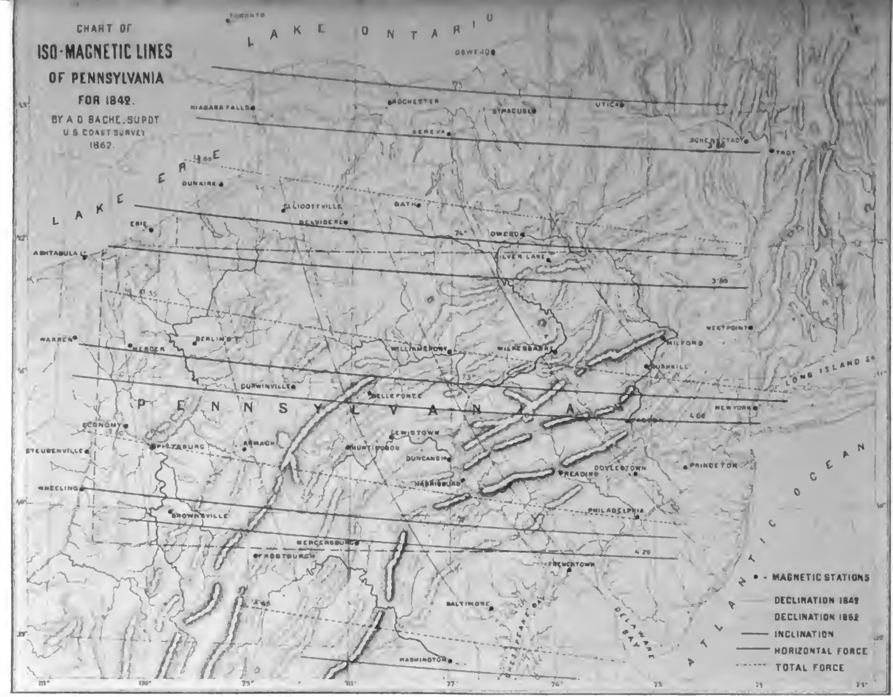


Fig. 6. Bache's chart of isomagnetic lines for 1842 for the state of Pennsylvania.

Geologist of Maryland. Still more recently, detailed magnetic surveys of North Carolina and Louisiana have been effected.

When made Superintendent of the United States Coast Survey in 1843, Bache ardently promoted a program of magnetic observations and discussions in that Bureau and, on the foundations laid by him, investigations of the Earth's magnetism were fostered and furthered by Schott. Schott, who for over 52 years (1848–1900) was connected with the Computing Division of the Survey, had much to do with the discussions of magnetic data and many of his numerous publications were devoted to that subject. The award of the Henry Wilde prize by the French Academy of Sciences only two years before his death was made on the ground that "the whole of his work furnishes one of the important contributions in the history of terrestrial magnetism." The geomagnetic work of our Government is a model one. It was reorganized under a much more intensive program in 1899 by L. A. Bauer, with the able assistance of D. L. Hazard under the superintendency of Dr. Pritchett, in the Section of Terrestrial Magnetism, now designated as the Division of Geomagnetism and Seismology. Bauer was succeeded as Chief of this Division by Faris, Braid, Hazard, Watkins, and Captain Heck, who since 1922 has been its progressive and enthusiastic leader. The Division is responsible for the continued magnetic survey of the United States and its territories and the operation of five widely distributed observatories in Maryland, Arizona, Alaska, Hawaii, and Puerto Rico.

During the last half of the nineteenth century, magnetic observatories were maintained by the Survey at Key West, Florida, from 1860 to 1865, at Madison, Wisconsin, from 1876 to 1880, at Los Angeles, California, from 1882 to 1889, and at San Antonio, Texas, from 1890 to 1895. Besides these the United States Government took part in two International Polar Years, the first in 1882–83 by establishing magnetic observatories at Point Barrow, Alaska, and at Fort Conger, Ellesmere Land, and the second in 1932–33 through the Coast and Geodetic Survey at College, Alaska, and the United States Weather Bureau at Point Barrow, Alaska, both of these with the cooperation of the Department of Terrestrial Magnetism of the Carnegie Institution of Washington.

The frequent shifting of the magnetic observatories of the Coast and Geodetic Survey in earlier years from place to place was necessary because of limited funds, but since the formation of the

Division of Terrestrial Magnetism in 1899 there have been in operation for nearly 40 years the five observatories listed above. To Canada, however, belongs the credit of having maintained in the Western Hemisphere a magnetic observatory for the longest period. The Toronto Observatory was established in 1840 by Riddell, who was succeeded by Lefroy under the direction of General Sabine to whose indefatigable energy we owe the splendid simultaneous program of work of the British colonial observatories.[1] The Toronto Observatory continued until the end of 1897, when it had to be transferred to Agincourt some miles east of Toronto because of electric-car disturbances. Thus Canada has the distinction of a continuous record of magnetic fluctuations for a century with the exception of a single year, namely, 1898.

Another eminent scholar of the period, Joseph Henry, was interested in the Earth's magnetism and in 1830–31 made a series of intensity-observations at Albany, New York. In his work he used a needle which had formerly belonged to Hansteen of Norway and another which had been used by Sabine. He observed a disturbance of the magnetism of the Earth in connection with the appearance of an aurora.

Besides those already mentioned, some of the other early pioneers in magnetic work in North America were Botanist David Douglas, who determined declination and horizontal intensity along the northwest coast of North America during 1829–34; Colonel Graham, topographic engineer of the United States, who observed magnetic intensity on the southern boundary of Canada during 1830–45; and Emory, who made observations on the Isthmus of Panama in 1849.

Magnetic observations in South America were relatively few until the last quarter of the nineteenth century. In 1819 Admirals Roussin and Givry made magnetic observations on the Brazilian coast between the mouths of the rivers Marañon and La Plata. In 1825 Boussingault observed in Ecuador at Marmato and Quito. In 1826–30 Captain Philip Parker King obtained data in southern portions of the eastern and western coasts of South America (Brazil, Uruguay, Straits of Magellan, and Chile). During 1831–36 Captain Fitzroy on the *Beagle* obtained data on the most southern portions of America. Sir Edward Belcher's voyage of 1837–42 secured magnetic data on the western coast of South America. During 1838–42 our own Wilkes Expedition secured many magnetic observations in the antarctic regions. In Brazil, during 1880–84, a Dutch Commission led by E. van Rijckevorsel, with the authorization and valuable assistance of the Brazilian Government, made an extensive survey of the eastern part of that country, including the valley of the San Francisco River. In 1910–11, chiefly for the purpose of determining secular variation in that valley, the Director of the National Observatory at Rio de Janeiro conducted a resurvey. The Government of Argentina has long maintained an excellent magnetic survey and three magnetic observatories (Pilar, La Quiaca, and Orcadas). Since its organization, the Carnegie Institution of Washington, through its Department of Terrestrial Magnetism, has occupied more than 2,000 stations on land in all the countries of South America and 1,000 at sea in the Western Hemisphere—a contribution of considerable importance in little-explored tropical regions of South America. The Huancayo Magnetic Observatory of that Department, established in 1919 with the approval of the Peruvian Government, in 12° south latitude, was then only a few miles south of the magnetic equator. Thus since Humboldt's observations of 1802 the magnetic equator has moved southward 5° in latitude and this southward progress is now continuing at the rate of about four miles per year. The force, too, has changed considerably since 1802 but is now nearly steady at about 0.30000 absolute unit, a decrease of 15 per cent in 140 years.

---

[1] Humboldt states that Celcius was the first to make use of observations at observatories to institute simultaneous measurements at two widely separated places.

The attack on problems of such broad scope requires a special type of research and, in unique degree, world-wide coordination of data and experiment. No single well-planned experiment or observation can solve the problems presented. Observations must be made in all parts of the world and must be continued over a long period. Techniques for the organization and interpretation of these data must be developed and experimental researches must be conducted along lines which will supply information on basic properties related to the subject. The Carnegie Institution of Washington, recognizing the extent of this field, established its Department of Terrestrial Magnetism in April 1904 "to investigate such problems of world-wide interest as relate to the magnetic and electric conditions of the Earth and its atmosphere, not specifically the subject of inquiry of any one country, but of

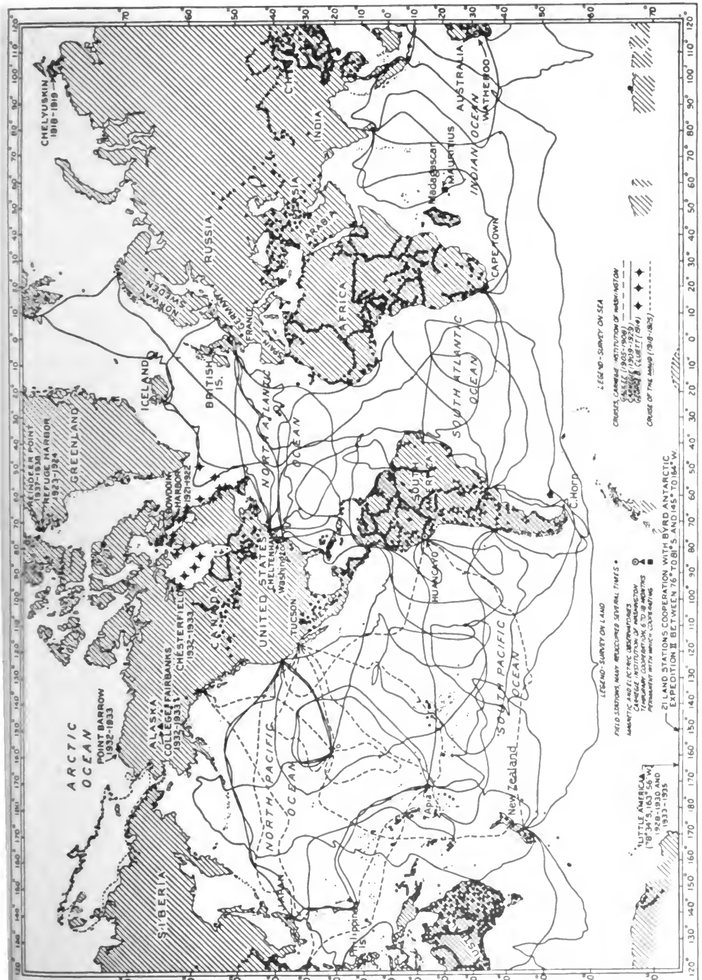


FIG. 7. World magnetic survey, Carnegie Institution of Washington, 1905–40.

international concern and benefit." Bauer, with the agreement of the Coast and Geodetic Survey, was called upon to be the organizer and first Director of this Department. Thanks to his enthusiasm and foresight, it has taken a leading part in obtaining data in magnetically unexplored lands and over the oceans, and in the determination and maintenance, in cooperation with the United States Coast and Geodetic Survey, of magnetic standards.

THE MAGNETIC FIELD AND ITS DETERMINATION

The Earth's magnetic field extends far out into space. Four thousand miles above us it is still one-eighth as great as at the surface. The Earth may be regarded as approximately a uniformly magnetized sphere with its axis making an angle of 11°.5 with the axis of rotation. Although but feebly magnetized as compared with the magnetization attainable in high-grade magnet steels, the

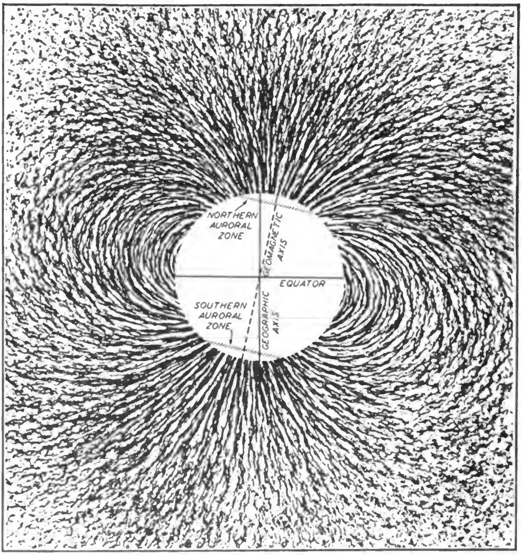


FIG. 8. The magnetic field about the Earth.

average intensity of magnetization per cubic centimeter for the entire Earth is roughly 10,000 to 100,000 times as great as the magnetization of ordinary crustal rocks. Such magnetization would be produced by the presence of one magnetized embroidery needle in every five cubic centimeters of the whole Globe. Appreciable irregularities in the field exist, but they do not cause great departures from the field which would be produced by the hypothetical uniform magnetization.

It is well known that the axis of the magnetic field is inclined to the geographic axis of the Earth and that the polar auroral zones are centered not on the actual magnetic poles but on the poles of uniform magnetization, which takes account of so great a part of the whole field. A general idea of the field, which, however, is by no means so simple, may be obtained by the distribution of iron filings over a disk magnet. The principal magnetic poles are distant 1200 miles or more from the geographic poles. The north magnetic pole, visited in 1831 by Ross and in 1903 by Amundsen, is in Boothia Peninsula in north Canada (latitude 70° 30′ north, longitude 95° 30′ west). The south magnetic pole, reached in 1909 by E. David, Douglas Mawson, and A. Mackay, of Shackleton's British Antarctic Expedition of 1907–09, is in South Victoria Land of the Antarctic Continent (latitude 72° 25′ south, longitude 155° 16′ east). Thus the line joining the magnetic poles is not a diameter of the Earth but passes at a distance of some 750 miles from its center.

Measurements to determine the Earth's magnetic field at any point must include observations of three magnetic elements, namely: (1) Magnetic declination, the angle between the true astronomical north-south meridional plane and the vertical plane through the magnetic north-south direction as defined by the compass; (2) magnetic inclination or dip, the angle through which a magnet entirely free to move would dip below the horizon in the magnetic north-south meridional plane; and (3) the total magnetic force, acting in the magnetic meridional plane or its horizontal component, or its vertical component. Painstaking and patient recording of the complex geomagnetic time-variations through years and decades, magnetic surveys on land and sea, and the reduction of all these data have disclosed certain systematic features. The scattered observations of declination run back over 300 years in a few limited localities, while those of inclination, owing of course to its smaller practical utility as well as to the greater difficulty attending its measurement, are much less abundant. It is now over 100 years since a method was first described to obtain the absolute measure of the intensity of the magnetic field. The classical methods of observation and discussion have reached a point where decisive progress demands the introduction of new experimental and theoretical ideas. Thus the present trend in geomagnetism is for greater collaboration with other branches of physics, geophysics, and solar physics.

There has been marked improvement both of instruments and accuracy of observation. The Bache magnetometer and the Robertson dip-circle used in Bache's magnetic survey of Pennsylvania were from the best instrument-shops of Europe. That observers were burdened in transportation during those early days of pioneer and difficult conditions of travel as compared with those of today may be readily realized when it is stated that the suspended magnet was 50 centimeters (20 inches) in length. The development of improved magnet steels has made possible instruments of far less weight and far greater accuracy with magnets one centimeter (less than one-half inch) in length. Thus a recent type of horizontal-intensity magnetometer developed in Denmark is so light that it is transported by mail and gives a precision of one part in 20,000.

Improvement in instruments for use on land has also followed improved electrical measurements, so that now high precision is attained by electromagnetic coil-instruments such as the sine-galvanometer. In this type the classical magnetic method is replaced by the electrical method. The auxiliary magnetic field with which the Earth's field is compared is no longer produced by means of magnets, but by means of electrical currents in carefully standardized coil-arrangements of accurately known geometrical dimensions. The limit of accuracy of the new methods is determined by that of measuring currents of the order of one-tenth of an ampere. The main advantage of the new methods (of which several variants have been developed) is their applicability to any component of the Earth's field. Electromagnetic instruments are now used for control of absolute magnetic standards.

The oceans comprise the greater part of the Earth's surface. The moving support provided by a vessel makes observations at sea difficult. Gimbal suspensions are used and, until recent years, instruments were of the magnetic type. Now the electromagnetic method is used, as, for example, the marine earth-inductor, which yields results of greater precision than older types of the sea dip-circle. In determining direction and magnetic intensity at sea, modern optical technique and improved magnet steels provide greater accuracy.

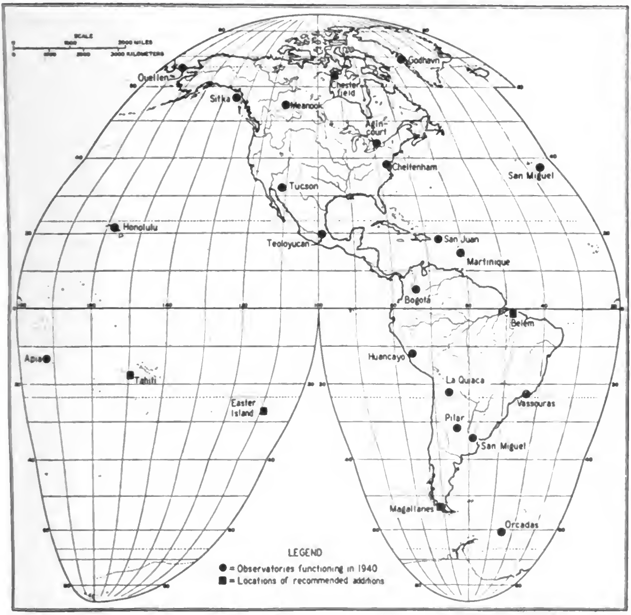


FIG. 9. **Established geomagnetic observatories of the western hemisphere, 1940, and recommended additions.**

The same holds true for equipment to record the multiple, regular, and irregular fluctuations of geomagnetism at observatories in the world net which now includes over 70 observatories. The older, massive variometers are replaced by small portable instruments and recorders.

The data supplied by observation on land and sea are the basis for charts of isomagnetic lines of declination or compass-direction, of inclination, and of intensity. The comparison of these charts, or of results at a single station for different epochs, shows slow changes of the elements—that is, secular variation. Thus we have historical records showing progressive shift of lines of zero magnetic declination (agonics) because of secular variation. On assembly of the data the shifts of isomagnetic lines are determined for successive epochs. We may then prepare an isoporic chart—that is, one showing lines along which the isomagnetics have made the same progress. Lines so drawn are called isopors, or lines of equal marching. The picture resulting from these charts presents the problem of secular variation in a new light and has raised questions

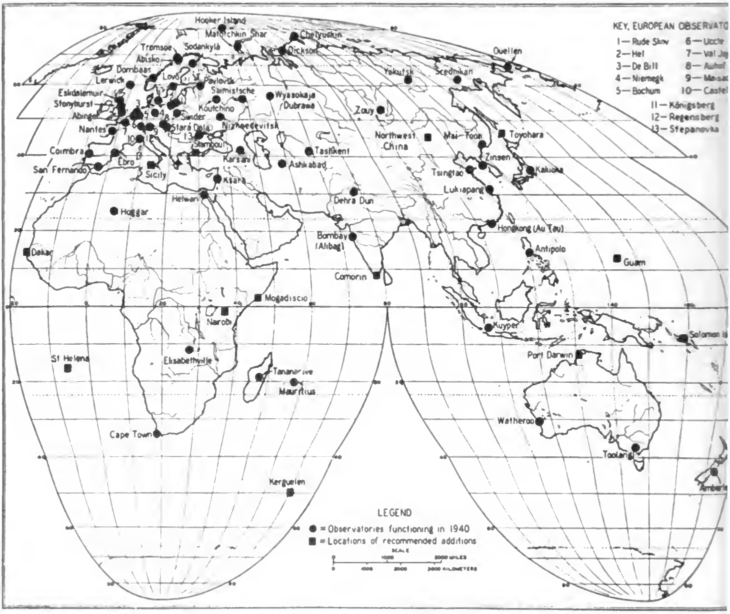


FIG. 10. Established geomagnetic observatories of the eastern hemisphere, 1940, and recommended additions.

of broad geophysical significance. The rise and fall of the rate of the secular change, the slow expansion, and then the gradual retraction of the areas within which there have been excessive alterations in any of the magnetic elements, indicate that changes are still active within the Earth's interior.

The maximum isoporic foci are practically all found in the hemisphere containing the great land-masses with the intervening Atlantic Ocean. Such foci as are found in the Pacific hemisphere are of but moderate intensity and not well defined. This relation of large and rapidly varying rates of change of the intensity and direction of the magnetic field to the surface structure of the Earth can scarcely be accidental; it is natural to suspect that there is a causal relationship existing between crustal or subcrustal movements and these magnetic manifestations.

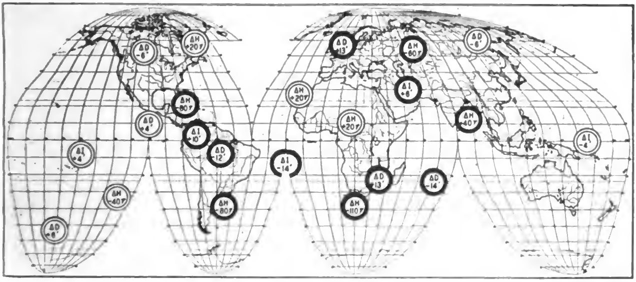


FIG. 11. World distribution of foci of rapid annual change of magnetic declination, inclination, and horizontal intensity, approximate epoch 1920–25.

The relatively short period of time during which we have observed, and may continue to observe, how the Earth's magnetic field changes, discourages hope of deriving an adequate understanding of the phenomenon by direct observation. This has led investigators to seek means for extending our knowledge of secular variation over a longer time-scale through the development of paleomagnetic research, that is, the determination of "fossil magnetization" of cores of material collected from the bottoms of the oceans and of varves deposited during glacial periods.

That irregularities of the distribution of the Earth's magnetic field over continental areas may bear definite relation to geologic features of the Earth's crust has long been recognized. It is only in comparatively recent years that anomalous magnetic data have been considered an aid in determining the character of geologic substructure. The difference between the uniform or normal magnetization and that actually existing serves to bring out the general irregularities as a residual field. The inhomogeneity of the Earth's outer crust or shell, perhaps 25 or more miles thick, gives rise to many regions of local magnetic disturbance. Some magnetic ore-deposits are large enough to cause local poles and other irregularities of intensity twofold or even threefold the normal value in their neighborhood. The experimental and theoretical progress made in applying magnetic methods to the determination of deposits in the Earth's crust has been much more satisfactory than that with regard to the main field. A major feature of physical and mineralogical research is offered by large systems of volcanic dykes in South Africa; these are magnetized opposite to the present direction of the Earth's vertical component. In this general region of South Africa the magnetic horizontal component of the Earth's field has decreased 16 per cent of its present value in a 30-year period! With improved knowledge of magnetic anomalies the question of interrelation with gravimetric anomalies and deep-focus earthquakes may be studied. Gravimetric and magnetic anomalies frequently occur in the same regions. Possibly these and the deep earthquakes have a common origin in current-systems in the inner Earth.

The mathematical analyses of several investigators, based on the latest observed world-data available, show that the major portion, about 95 per cent of the Earth's total magnetic field at any one time, arises from causes beneath the Earth's surface and that two minor portions total about five per cent, of which one arises from causes above the Earth and one from vertical electric currents passing through the Earth's crust.

Thus the physical origin of the interior and principal part of the Earth's field still offers a baffling problem. Besides the study of secular variation over previously unattainable periods of thousands of years by utilization of the magnetic polarization of varves and deposits, we must look to other methods in the laboratory. One of the most promising of these is the investigation of materials under conditions of extremely high pressures and temperatures which approach those at great depths below the Earth's surface.

Cosmical Relations

It is the external part of the Earth's field represented by the primary geomagnetic time-variations, such as diurnal variation and magnetic perturbations, to which we must look for study of cosmical interrelations, and in these spectacular progress has been made. From the beginning of systematized magnetic observations and records there has been evident a general correlation between the Sun and the Earth's field. Thus with the discovery by Graham at London in 1722 of the variation during the solar day,[2] early investigators surmised such a relation. In 1741, Celsius and Hiörter at Upsala confirmed by a long series the connection between auroral displays and disturbances in the normal course of the needle. With the accumulation of observational data on the normal and disturbed conditions of the Earth's field from month to month and from year to year, it has long been apparent that a marked relation exists between annual magnetic means and the annual averages of relative sunspot-numbers.

This relation has been further disclosed by study of the data on the magnetic activity of the Earth obtained by many observatories (during 1938, 50 in number) by designating magnetic activity on an internationally adopted scheme of classification maintained since 1906[3] through the International Meteorological Organization and the International Association of Terrestrial Magnetism and Electricity on a scale of 0, 1, and 2. The mean values of such character-numbers—that is, intrinsically measures of magnetic activity for the whole Earth—have proved that there is an interrelation between conditions on the Earth and on the Sun through indication of a 27-day recurrence-tendency of geomagnetic activity, which is in harmony with the period of rotation of the central part of the Sun. Because of the largely qualitative type of this simple measure, there was adopted in September 1939, during the Seventh Triennial Assembly of the International Union of Geodesy and Geophysics, a new scheme for a quantitative measure of magnetic activity, namely, the magnetic three-hour-range index, *K*, varying from 0 to 9. Through joint studies of the United States Coast and Geodetic Survey and the Department of Terrestrial Magnetism of the Carnegie Institution of Washington, following some earlier experimentation at the Adolf Schmidt Magnetic Observatory at Niemegk, the three-hour-range index was established as a valuable abstract of the magnetograms, providing even single observatories with good estimates of world-wide magnetic conditions. This index is currently derived from data obtained by seven American-operated observatories [Cheltenham (Maryland), Tucson (Arizona), Sitka (Alaska), Honolulu (Hawaii), San Juan (Puerto Rico), Huancayo (Peru), and Watheroo (Western Australia)], and is published weekly in the United States by *Science Service*. It gives, for the first time, a detailed homogeneous series for the intensity

---

[2] Father Tachart, in the presence of the King, at Louveau, Siam, made hourly experiments, in 1682, which showed the compass-direction to be different on seven different days—probably these were made at different times of the day, and they were thus the first real indications of the diurnal variation.

[3] Compilations of this measure as obtained from a varying number of observatories have recently been made from 1905 back to 1884 by G. van Dijk.

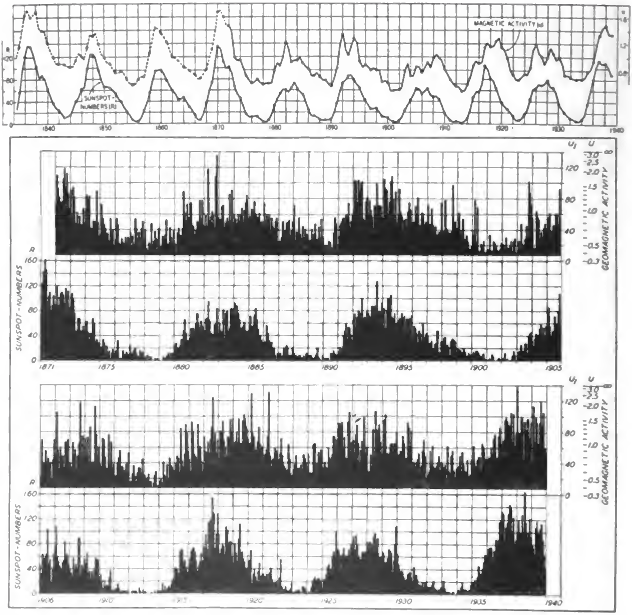


FIG. 12. Geomagnetic activity and sunspot-numbers: (*A*) average annual values for the past century; (*B*) monthly means 1872–1939.

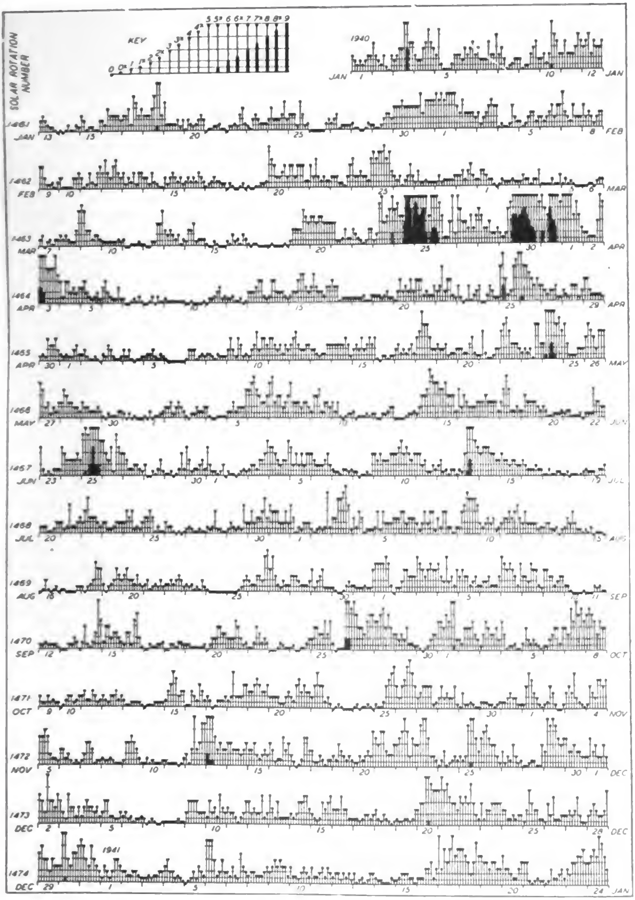


Fig. 13. Three-hour-range indices of geomagnetic activity, January 1, 1940 to January 24, 1941, plotted in 27-day periods according to solar rotations.

of solar corpuscular radiation affecting the Earth, useful both in its terrestrial aspect—as in scientific or commercial radio work—and for its bearing on solar physics. Violent magnetic storms with three-hour-range index of 9 occur only a few times near a sunspot-maximum, but it is equally rare that any full three-hour interval is perfectly free from disturbance. This means that the Earth is almost constantly, even near sunspot-minimum, under the influence of particles (presumably solar), weak as this influence may be at times.

The intensity of the ionizing solar wave-radiation absorbed in the ionosphere on the daylight hemisphere can likewise be measured geomagnetically in the amplitudes of the solar daily magnetic variation. Analysis of 18 years' records (1922–1939) of horizontal intensity at the Huancayo Magnetic Observatory, in which the magnitude of the solar daily magnetic variation is exceptional, gave a measure for the ionizing solar wave-radiation comparable with the relative sunspot-numbers as the only available complete series of daily measures of solar activity. The correlation-coefficient between them is + 0.92 for monthly means and + 0.984 for annual means. These are the closest relations so far established between phenomena on the Sun and the Earth.

Besides the solar diurnal variation, there is another which is small and quite systematic, namely, that associated with the Moon's position. The most conspicuous lunar effects so far found in any geophysical phenomenon—apart from the tidal observations of the sea—appear in the diurnal variation of magnetic horizontal intensity at Huancayo. These results can be interpreted by the dynamo-theory as the magnetic field of electric currents induced in the oscillating ionosphere by the Earth's permanent magnetic field. These currents must flow low in the ionosphere, where recombination of ions is rapid, because the lunar variation is found to be confined to the daytime; neither primary ionospheric currents nor secondary currents induced within the Earth's body cross the night hemisphere near the equator. The ionospheric air-motions appear to be much more dominated by tidal oscillations than are the winds near the ground, because the Moon's geomagnetic effect is plainly expressed in the magnetograms even for single days. The partial tides connected with the Moon's varying distance from the Earth are clearly recognized in the geomagnetic records, but the relations of the amplitudes and phase-angles of

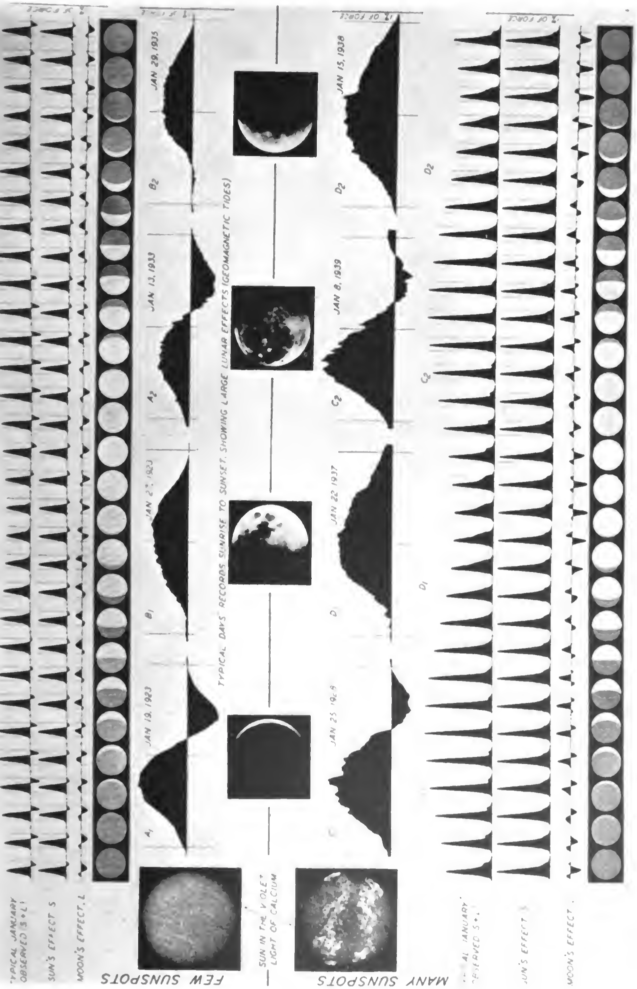


FIG. 14. Daily changes of geomagnetic horizontal intensity caused by Sun and Moon during sunspot-minimum and maximum, Huancayo Magnetic Observatory, Peru.

the geomagnetic partial tides to the main semidiurnal tide differ significantly from those in the gravitational tidal forces. The study of geomagnetic tides provides thus a new approach to the study of resonance-phenomena in atmospheric oscillations.

Other evidences of cosmical influences on the magnetic field of the Earth are the spectacular displays of auroral lights which are most intense during periods of perturbation of the Earth's field when they are always seen in polar regions. To explain these phenomena of many types, mathematical and experimental investigations have been made, especially by our Norwegian colleagues. I need hardly call to mind Birkeland's well-known experiment with a cathode sphere in a vacuum-chamber, by which he demonstrated the accumulation of electrons in the plane of the magnetic equator. Since that time his successors and distinguished colleagues Störmer and Vegard have done much to advance the study of auroral phenomena. Depths of penetration of the auroras are determined through simultaneous parallactic photographs and physical properties through interpretation of auroral spectra. It is found that polar-light beams generally do not come closer to the Earth's surface than about 50 miles; some come no closer than 300 miles or more. Elaborate calculations of the paths of electrified corpuscles which are entrapped between outermost lines of magnetic force of the Earth's field have been beautifully demonstrated in the laboratory by Brüche.

The regular daily magnetic variation is the effect of radiations from the Sun; these radiations travel with the speed of light and are absorbed in the highest layers of the atmosphere, making these layers electrically conducting. More than 50 years ago Balfour Stewart and Arthur Schuster inferred from geomagnetic data that there must be a high atmospheric layer of great electrical conductivity. This has been confirmed and extended by direct investigations of the ionosphere which were initiated at the Department of Terrestrial Magnetism in 1925. This fact became more important after the invention of wireless telegraphy because reflections from these conducting layers are found to be responsible for radio transmission. A striking confirmation has now been furnished by the recent development of the so-called ionospheric apparatus which produces records automatically by photographing echoes of vertically directed variable-frequency radio waves; from these records the electric characteristics of the several conducting

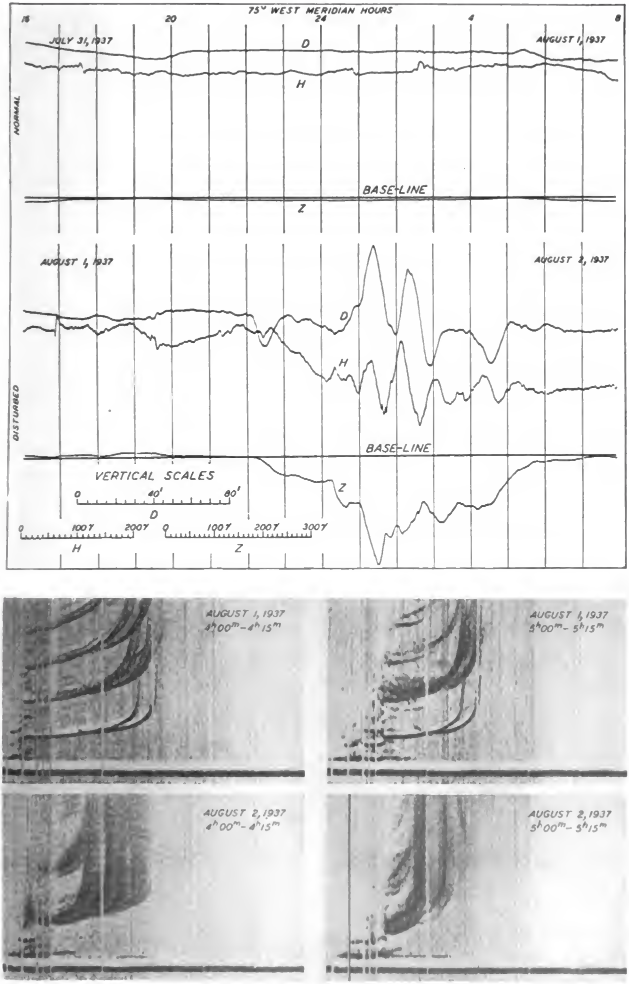


FIG. 15. Geomagnetic and ionospheric records showing magnetically quiet (July 31 to August 1, 1937) and moderately disturbed conditions (August 1 to August 2, 1937). [The ionospheric records for the two normal quarter-hours on August 1 are contrasted with the disturbed conditions resulting from the storm for the same quarter-hours on August 2.]

layers of the ionosphere may be deduced. There are frequent sudden fade-outs of high-frequency radio signals. Examination of geomagnetic records indicated these fade-outs to be associated with sudden disturbances in the Earth's magnetic field. This directed attention to the Sun's condition as a possible cause and astronomers were called upon to seek possible associations with solar phenomena. The astronomers were not long in finding that many fade-outs had occurred simultaneously with bright eruptions in the solar chromosphere.

Thanks to the extensive network of stations all over the Earth keeping constant watch for these chromospheric eruptions, a similar and older network of magnetic observatories, and the channels of radio communication, data on these associated phenomena accumulated rapidly. Dr. Dellinger's compilation of such data soon showed that chromospheric eruptions always accompany fade-outs and the concomitant magnetic changes. The first actual simultaneous records of a fade-out, a bright eruption in the solar chromosphere, and a particular bay-like kind of geomagnetic disturbance were obtained August 28, 1937. Characteristically the echoes disappear quite suddenly because of strong absorptions of the radio waves as they pass through the regions at a level between 70 and 90 km. Absorption is produced by the intense ionization of this region by the ultra-violet light emanating from the solar eruption.

The magnetic storm of March 24, 1940, probably the greatest magnetic storm ever recorded, was an event of unusual geophysical interest. In addition to the disruption of wire-communication by electric currents induced in the Earth, this storm induced earth-currents of such magnitude that electric power-systems were severely affected—the first time such effects have ever been reported. Computations were made of the intensity of currents which could be produced by magnetic changes, and for extreme cases it was found to be sufficient to produce the observed effects on power-lines. Thus an increased practical importance of cosmical research in geomagnetism has been shown, for the lengthy observations extending over a century supply definite information on the probability of occurrence of such storms and therefore on the extent to which it is advisable to improve electrical installations to avert their effects.

Our knowledge of magnetic storms is being rapidly increased

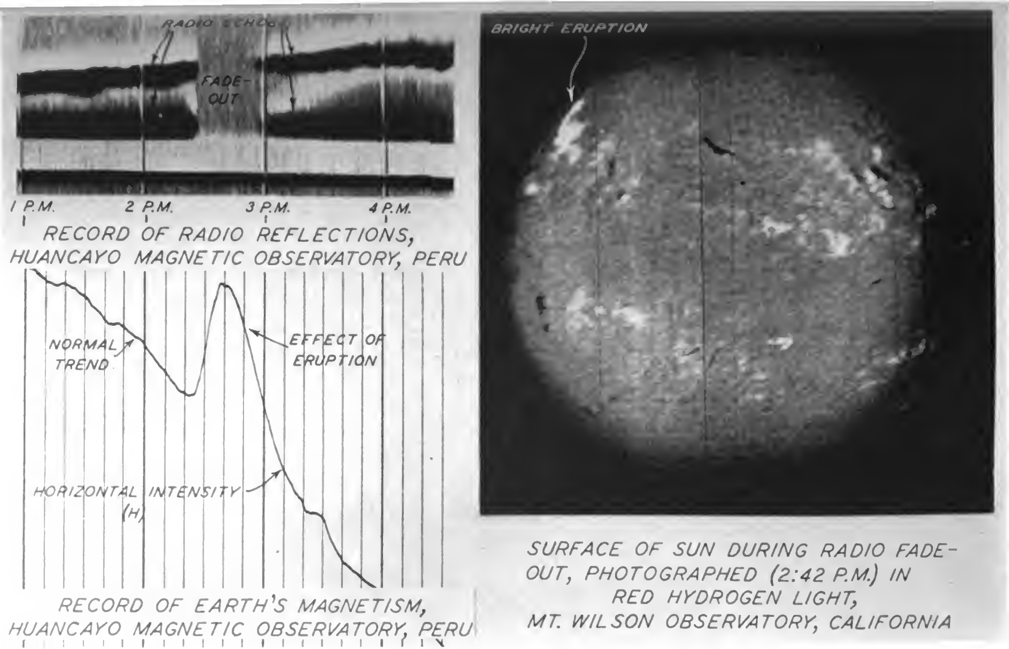


FIG. 16. Changes in Earth's magnetism and cessation of radio reflections from ionosphere accompanying bright eruptions in solar chromosphere August 28, 1937.

through study of the large body of data which has been made available through one of the most outstanding examples of co-operative research—the Second International Polar Year. Recognition that facilities must be provided for securing records of the great magnetic storms has caused a number of observatories to install insensitive magnetographs capable of obtaining complete tracings for greater storms than have ever been recorded since regular magnetic recordings were begun nearly a century ago. With the aid of these insensitive magnetographs, the very great storms of April 1938 and of March 1940 were completely recorded at a number of observatories.

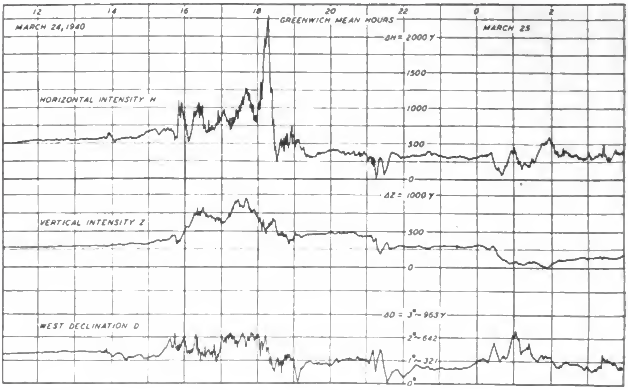


FIG. 17. Geomagnetic records obtained at Potsdam-Niemegk, Germany, during the great magnetic storm of March 24–25, 1940, the greatest storm of which there is record.

Another connection between geomagnetic and cosmic phenomena has only recently been discovered—the world-wide decrease in the intensity of cosmic radiation during great magnetic storms. Although such an effect was predicted some years ago, observational verification was not achieved until the occurrence of the great magnetic storms of the present sunspot-cycle. The effect was noted by many observers during the great storms of April 1937 and of March 1940. Since the magnetic moment of the Earth increases during the main phase of a magnetic storm, a

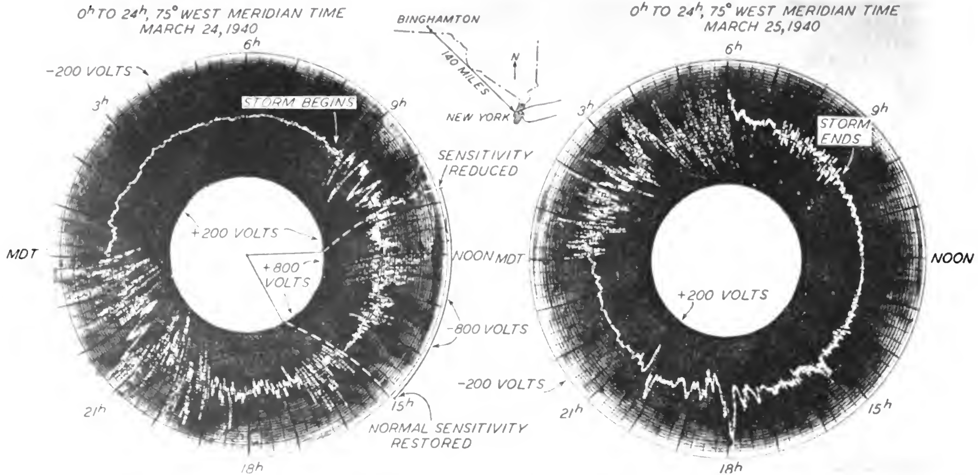


FIG. 18. Records showing changes in earth-potential on Western Union telegraph line New York, New York, to Binghamton, New York, during the great geomagnetic storm of March 24–25, 1940.

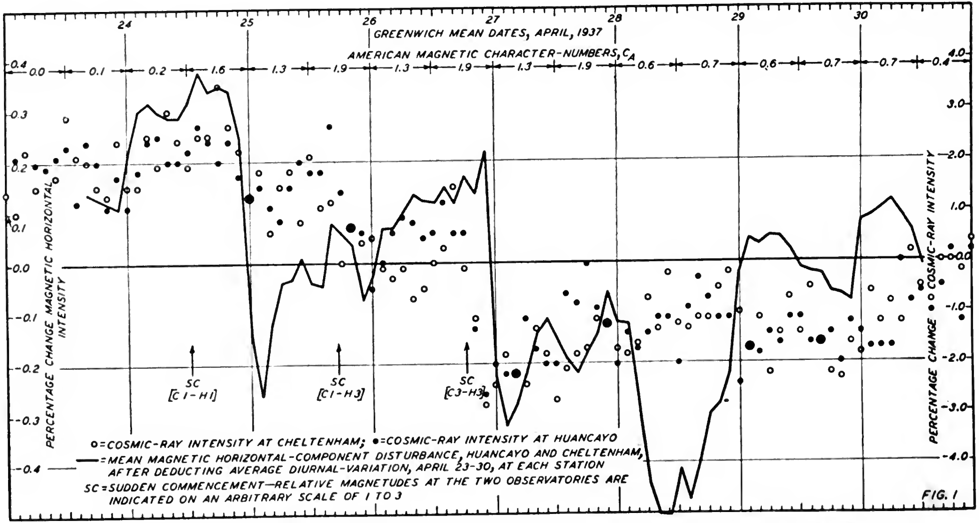


FIG. 19. Bihourly departures expressed in percentage of absolute values for cosmic-ray intensity and for geomagnetic disturbance April 23 to 30, 1937.

decrease of cosmic radiation during that time is to be expected on the basis of the well-established variation of cosmic-ray intensity with geomagnetic latitude. However, the magnitudes of the expected and of the observed effects differ greatly. How may this discrepancy be reconciled, and what bearing does it have on our notions regarding magnetic storms? Some have thought that if the main phase of the storm-time field is due to a ring-current, the concomitant change in cosmic-ray intensity may be reconciled if we regard the radius of the ring-current to be very great. But the trajectories of cosmic-ray particles in the presence of a large ring-current are exceedingly complicated, and there is reasonable doubt whether the observed effects could be produced in this manner. Mathematical treatment has been unsuccessful in dealing with the problem. It has also been found that the cosmic-ray intensity is practically always lower for the five international magnetically disturbed days than for the five international magnetically quiet days of each month. This indicates that, on the average, the cosmic-ray intensity is lowered on days of magnetic disturbance, which is in accord with the magnetic-storm effect. It has also been discovered that the 13½-day and 27-day waves in cosmic-ray intensity are closely associated with those for character-figure and magnetic horizontal intensity. This directs attention to a possible means by which cosmic-ray data might serve to establish the existence of a general solar magnetic field.

This brings us to another promising attack in the laboratory, namely, the attempt to enlarge knowledge of the basic nature of magnetism through investigation of the detailed laws governing the interaction of the magnetic particles composing all matter. All the actions and reactions in the world of physical things may be expressions of three fundamental forces—gravitational, electromagnetic, and nuclear. Perhaps all three may ultimately be reduced to different aspects of the same all-pervading, all-inclusive type. Perhaps some obscure atomic effect of the extremely high pressures of the centers of such bodies as the Earth and the Sun may be the cause of their magnetization. Therefore, equipment has been recently devised for the experimental studies of nuclear physics through the use of high voltages such as the electrostatic generator and the cyclotron. The Carnegie Institution of Washington has, since 1938, had an electrostatic generator capable of developing some five million volts and has now partly completed a cyclotron which may reach 15 million volts. The electrostatic generator operates under 50 to 60 pounds gauge-pressure and is enclosed in a large steel tank 55 feet high in the so-called Atomic-Physics Observatory—"Observatory" since we look at the infinitely small as our astronomical colleagues look at the infinitely large in space.

CONCLUSION

Looking back over three centuries of research, we see that the general interest of scientists in the Earth's magnetism has been waxing and waning in the course of time. In the days of Gilbert and Halley, and in the classical period of research a century ago, when Bache and his colleagues made their contributions from America, the Earth's magnetism was studied in the belief that it might prove to be a fundamental phenomenon similar to the Earth's gravity. When these hopes were not easily fulfilled, the interest turned to other branches of science seemingly abundant in far-reaching discoveries. Because of the significant recent advances outlined in this address, however, geomagnetism has again obtained a key position as a thriving branch of geophysics, not only because of its own intrinsic interest, but mainly because geomagnetic conditions represent a complete, a faithful, and an intelligible record of cosmical influences to which the Earth and all human beings are subjected. We may well agree with Humboldt that "Nothing that occurs in our planet can be supposed to be independent of cosmical relations."

AURORA AND GEOMAGNETISM

C. W. GARTLEIN

Curator, Department of Physics, Cornell University

(Read February 15, 1941, in Symposium on Geomagnetism)

THE polar aurora (borealis and australis) is a luminosity of the night sky which may be seen at certain times from almost any point on the earth, though only rarely in the Tropics. It is characterized by certain forms, colors and variations. The aurora is usually yellowish green in color, but red and violet often appear. The forms may be put into four general classes: glows or diffuse surfaces; arcs and bands; rays, draperies and coronas; and pulsating forms. Most of these have special significance for geomagnetism. The work of Störmer and others has shown that these displays occur in the earth's atmosphere at heights 60 to 1,000 km. The spectrograph has revealed that the light is radiation from the usual atmospheric constituents, oxygen and nitrogen, which have been excited in a special manner, presumably by high speed particles.

Serious study of the relations of the aurora to geomagnetism began in 1741, when Celsius and Hiörter pointed out that magnetic disturbances and auroras occurred at the same time at Uppsala. Graham, at London, cooperated with Celsius and found that magnetic disturbances occurred at the same time there as at Uppsala. Sometimes the magnetic disturbance preceded the aurora. Sabine and associates showed that great magnetic storms are accompanied by auroras, but small disturbances are not. Conversely, in the Arctic regions auroras often occur without much magnetic disturbance, that is, change in declination. At the close of the nineteenth century Angot remarked that in middle latitudes auroras and magnetic storms seemed closely related, but in arctic regions the relation was not closely followed. He suggested that arctic auroras might be different from temperate auroras. We will consider these questions in more detail after we look at other early geomagnetic and auroral recordings.

It was early noted that auroral arcs stretch across the sky

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 2, MAY, 1941

almost at right angles to the magnetic meridian. This observation is not wholly supported by later work. The rays of the aurora were known to be parallel to the freely suspended compass needle and coronas appeared over the magnetic zenith as first noted in 1747, and the corona moved as the magnetic inclination changed.

To summarize our present knowledge, we may say that the earth's magnetic field varies with latitude and longitude. It has secular variations in intensity, declination and storminess as well as diurnal and lunar variations. The magnetic storms follow an eleven year cycle about in step with the sunspot cycle. These storms have a tendency to recur at twenty-seven day intervals and are somewhat more frequent at the equinoxes. Earth currents occur during these storms. For each of these statements there is one closely similar for the aurora borealis.

It is now common knowledge that auroral displays are most frequent in polar regions. Fritz, in 1881, published a map showing this frequency variation, which was first suggested by E. Loomis in 1869. Since the data used in deriving the map were necessarily taken by many observers, the curves are not as precise as desired but still give us the best information we have. We learn that only one aurora is seen, on the average, each year in Florida while about twenty are seen in Maine, and about one hundred per year over Hudson Bay. However, only about one-third as many displays are seen each year in western United States as in the same latitude in eastern United States. About five times as many displays are seen at a given latitude in the United States as are seen at the same latitude in Europe. These lines of equal number are plotted as ovals whose approximate center is close to the geomagnetic axis pole and the curves themselves approximate geomagnetic parallels. The auroral zone is usually defined as a region containing the curve of maximum frequency, about 100 per year. This zone is of about 23° radius around a point in latitude 78° N., longitude about 68° W., the north geomagnetic pole.

The number of auroras seen per year in New York state and nearby is plotted in Fig. 1. It should be noted that during a long period at least twenty-five auroras were seen each year and as many as one hundred and ten in one year.

Other correlations with the earth's permanent field are found in the positions of auroral arcs and the angles of inclination the auroral rays make with the earth's surface. Observation in northern Europe in the nineteenth century indicated that the highest point of an auroral arc was on the magnetic meridian and the arc was perpendicular to it. Observation at other points, especially in arctic regions, showed this was not strictly true. The work of Störmer and Vegard has shown that for regions inside the auroral zone the direction of arcs is quite variable but for stations outside the maximum zone arcs lie nearly along the geomagnetic parallels with the western end 5° to 10° north of the parallel. The work of Vestine and Chapman shows that arcs are about parallel to the calculated direction in which currents would flow in the upper atmosphere to produce certain magnetic disturbances.

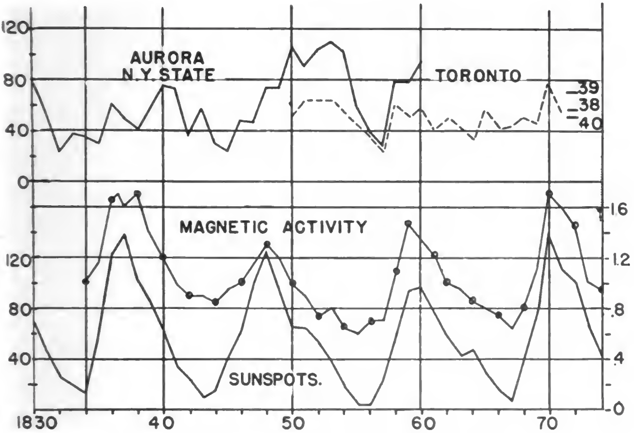


Fig. 1. Auroras, magnetic activity and sunspots.

The early observation, that auroral coronas (Fig. 2) are formed at the magnetic zenith (opposite to direction of magnetic dip), has been universally confirmed. Vegard has summarized many observations and concluded that for all stations the auroral corona is at the magnetic zenith of that particular time. The magnetic zenith is shifted a little lower in the sky during auroral displays, which are usually times of magnetic disturbance.

The corona is, of course, a perspective effect, as rays are everywhere parallel, as can be seen by looking north, or sometimes south as during the great aurora of August 11th–12th, 1939. Figure 3 shows a photograph of a bundle of rays appearing toward the east from Ithaca, on October 15, 1939. These rays make an angle of about 70° with the horizon while the quiet day dip angle is about 73°. The rays are toward the eastern end of a faint horseshoe curved band. Faint pulsating patches also were recorded by this photograph.

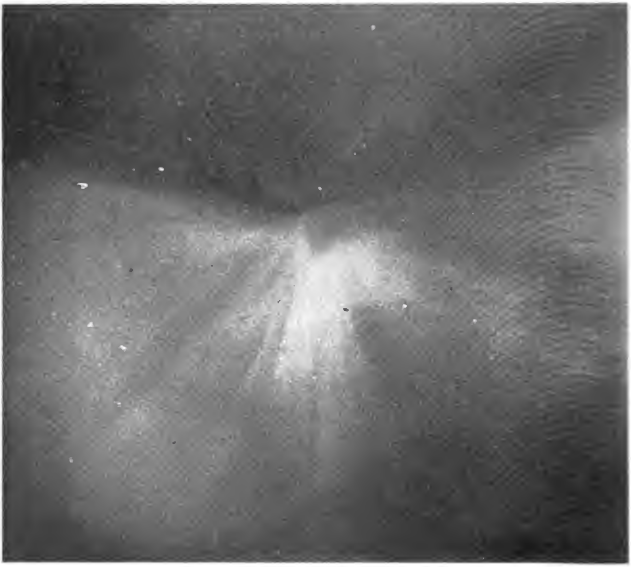


Fig. 2. Auroral corona at Ithaca, April 17, 1939, 1:17:52 A.M., within 2° of normal magnetic zenith.

The aurora appears at irregular intervals but when the number of appearances in each year is plotted over many years (Fig. 1) we find a rise and fall in eleven year cycles, though with some lag from the peak of the sunspot cycle. The magnetic activity curve follows the sunspot cycle more closely in time, the lag being

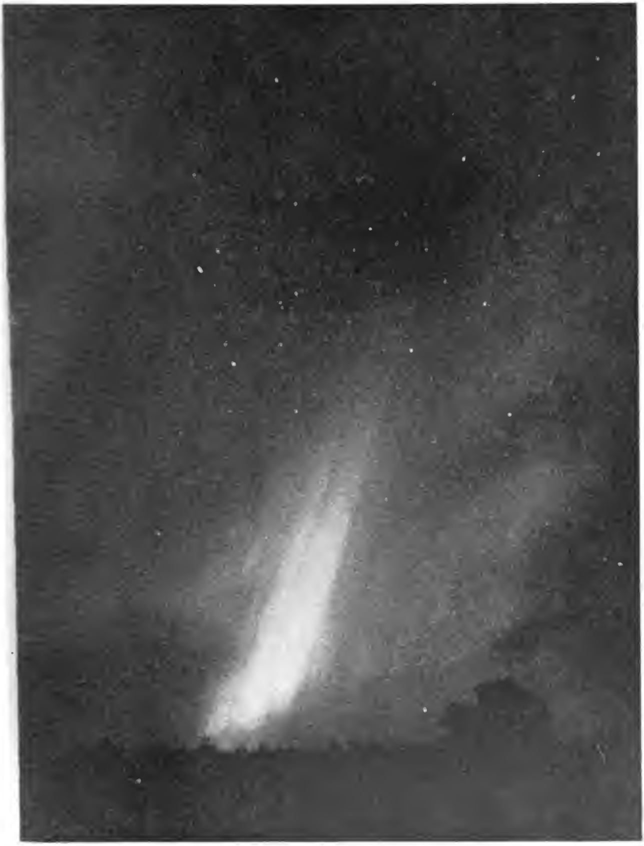


Fig. 3. Rays east of Ithaca, October 15, 1939, illustrate parallelism with earth's field.

about six months. A solar influence on both auroras and magnetic activity is indicated. The three phenomena also exhibit a long time secular variation of unknown period. A diurnal period is known in geomagnetism but not clearly proved in auroral work except for certain stations in arctic regions where maximum activity occurs at about eleven P.M. by magnetic time. A small effect of the moon on geomagnetism is known but no similar influence is shown on the aurora except that moonlight outshines the weaker displays whose presence can then only be detected by spectroscopes.

Figure 4 shows the total number of auroras seen in New York state in each month during the period 1830–1870. The aurora is thus most frequent near the equinoxes as are magnetic storms. Auroras have a tendency to recur after twenty-seven days as do magnetic storms. We have studied this tendency for some of our data and find that the recurrence after twenty-seven days occurs about as often as occurrences on two successive nights. The first occurs one time in five and the second about one in four.

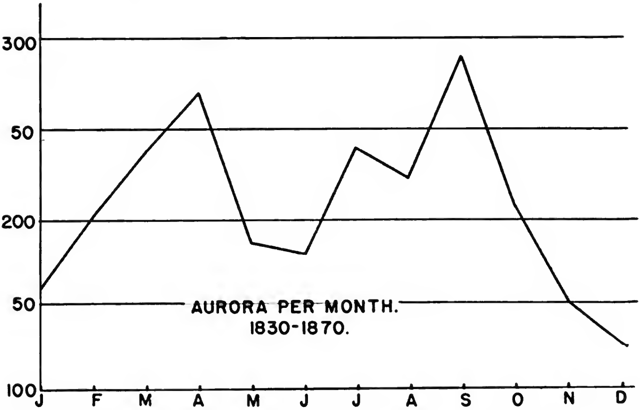


FIG. 4. Total auroras in each month for the years 1830–1870 corrected to 31 day months.

We have several times above noted a correlation between magnetic "storms" and auroras. (A magnetic storm is a somewhat erratic fluctuation of the magnetic elements over a period which ranges from a few hours to several days. The magnetic intensity changes only a few per cent.) This relation has been noted in many cases but data from different sources are conflicting.

It is generally agreed that in lower latitudes (geomagnetic 50°), such as that of Washington, auroral displays accompany intense magnetic storms. It was formerly believed that magnetic storms and auroras in polar regions were not always related. Vegard showed that strong auroras often occur at Bossekop, Norway, in the early evening during only minor magnetic disturbance. The work of the Byrd expedition showed that increased auroral activity occurred on the same day or the day following a maximum in magnetic activity. Whenever close correlations in time between these phenomena are considered we find discrepancies as though one were not governed only by the other. Little data were readily available on auroras in the United States so the National Geographic Society and Cornell University began a study in 1938. Photographic stations were set up at Ithaca, Geneva and Hamilton, New York, and many amateur astronomers have aided by making visual observations. Enough data will soon be available to study some of these correlations in detail.

Using these data taken at Ithaca and a few other stations during the interval March, 1938 to October 14, 1939, we have drawn Fig. 5. We have selected the nights of aurora on which we had sufficient data to assign a character figure for each three hour period of observation. The auroral characters are defined approximately as follows: 0 = no aurora; G = a small glow; 1 = an arc or weak rays; 2 = a display of rays extending up 40° from the horizon; 3 = a bright display extending nearly to the zenith; and 4 = a brilliant display, usually extending south of the zenith. The magnetic K figures indicate the maximum range in variation of the most disturbed element during the three hour period as defined in papers by Bartels, Heck and Johnson. During the one hundred twenty-three nights of aurora we had fifteen periods of auroral character 0, eighty-eight of G, thirty-eight of 1, forty-four of 2, thirty of 3, and eleven of 4. The chart was prepared by tabulating the number of times in the above group a given auroral figure, say 2, occurred when the K figure was a certain value, say 5. This particular case happened eleven times. Similarly, for other values. We find a definite correlation, the coefficient of correlation being .63. The lack of one hundred per cent correlation is not due alone to the fact that auroral numbers are partly estimated, but appears to indicate the effect of another agency, possibly the state of the higher atmosphere into which the disturbance producing particles come. A similar correlation but using the magnetic number from the previous three hour period gave a correlation of only .49, and a much poorer correlation between high magnetic numbers and high auroral numbers. Further study will shortly be made.

It is known that large earth currents flow during times of magnetic storms and correlations between these phenomena are very good. Similarly a comparison of earth current data (obtained by the Bell Telephone Laboratories) with auroral records indicates that in many cases the greatest currents flow within a few minutes of the peak of the aurora and all nights of aurora are periods of earth current activity.

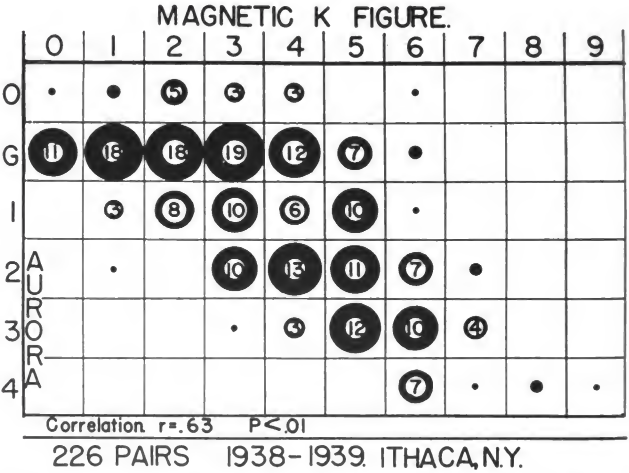


FIG. 5. Correlation between auroral and magnetic character figures for the same 3 hour intervals.

In conclusion we may note that while correlations between geomagnetism and aurora are not usually of an obvious cause and effect nature, it seems necessary that any theory of the magnetic storms must also account for the auroral displays.

REFERENCES

J. A. **Fleming** (editor), *Terrestrial Magnetism and Electricity*, McGraw-Hill, New York, 1939.

E. W. **Hewson**, Survey of Facts and Theories of the Aurora, *Rev. of Mod. Phys.*, **9**, 403–431 (1937).

Bartels, Heck and Johnson, Geomagnetic Three-hour Range Indices for Years 1938 and 1939, *Terr. Mag.*, **45**, 309–337 (1940).

Vestine and Chapman, The Electric Current System of Geomagnetic Disturbance, *Terr. Mag.*, **43**, 351–382 (1938).

F. B. **Hough**, *Results of Meteorological Observations in the State of New York 1826–1850* Albany, 1855.

F. B. **Hough**, Same, Second Series, Albany, 1872.

CONTRIBUTIONS OF IONOSPHERIC RESEARCH TO GEOMAGNETISM

L. V. BERKNER

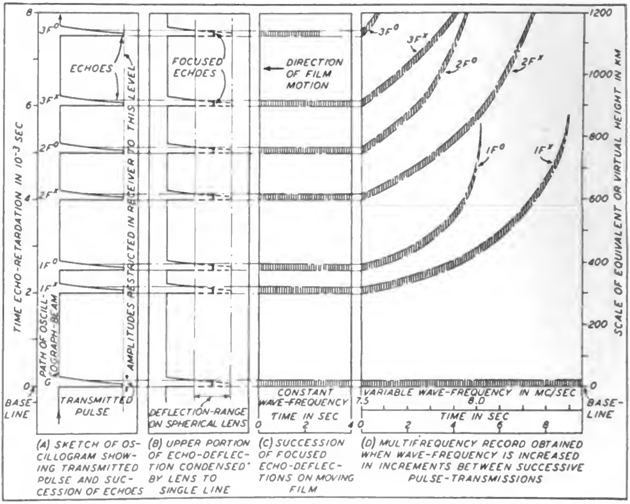
Department of Terrestrial Magnetism, Carnegie Institution of Washington

(Read February 15, 1941, in Symposium on Geomagnetism)

The concept of an electrified outer atmosphere as a source of geomagnetic and auroral effects dates back to Professor Balfour Stewart in 1882, though the earlier writings of Gauss consider its possibility. Subsequently, the names of Schuster, Kennelly, Heaviside, Lorentz, Eccles, Larmor and many others, loom large in the development of the concept.

Ionized layers of the outer atmosphere were observed directly by Appleton, and by Breit and Tuve in 1925 by somewhat different methods. Both used radio waves. Of these, the echo-method of Breit and Tuve has now been adopted almost universally. By this method a short pulse of radio waves is sent upward and the time measured for the echo to return. An equivalent or virtual height to which the wave would travel at this velocity can be computed assuming it to have traveled at the velocity of light. The density of ionization which reflects the wave at this level is measured by the wave-frequency transmitted. Corresponding to each wave-frequency is a particular density of ions which will reflect the wave. The virtual heights of each of all ion-densities can be measured, as in Fig. 1-*D*, when measurements are made on each of all necessary wave-frequencies. Such a series of determinations can be formed into a record such as shown in Figs. 2-*A* and *B*. From data of this kind, the actual heights and densities of the ion-layers of the outer atmosphere are determined. This determination depends upon the "magneto-ionic theory."

Complete experimental confirmation of the magneto-ionic theory constitutes one of the triumphs of ionospheric research. This theory provides the tool for exact measurement of equivalent electron-densities at various levels using radio waves. Let us consider some of the confirming evidence. For instance, the theory predicts that a radio wave, propagated in the ionosphere in the presence of the Earth's magnetic field, will be split into two wave-components. The behavior of these wave-components will be more complex than the simple wave transmitted and they will be propagated differently; one wave-component will be reflected at a lower electron-density than the other. Observation shows that two such components are returned from the ionosphere. The theory says that the wave-frequency at which each wave-component will penetrate a layer will differ by an amount which is a function of the strength of the geomagnetic field. Observation shows this separation of penetration-frequencies to be just the amount predicted. At Washington the difference is about twice that at Huancayo where the geomagnetic field-strength is only half as great. Finally, the theory states that at the magnetic equator, where the field is horizontal overhead, the two reflected wave-components will vibrate in mutually perpendicular planes, one always along the field, the other perpendicular to it. At


TIME ECHO-RETARDATION IN 10^-3 SEC
8
6
4
2
0
BASE-LINE
$3F^o$
$3F^x$
$2F^o$
$2F^x$
$1F^o$
$1F^x$
G
ECHOES
PATH OF OSCILLOGRAPH-BEAM
AMPLITUDES RESTRICTED IN RECEIVER TO THIS LEVEL
TRANSMITTED PULSE
FOCUSED ECHOES
DEFLECTION-RANGE ON SPHERICAL LENS
DIRECTION OF FILM MOTION
CONSTANT WAVE-FREQUENCY
TIME IN SEC
0
2
4
VARIABLE WAVE-FREQUENCY IN MC/SEC
7.5
8.0
TIME IN SEC
0
2
4
6
8
SCALE OF EQUIVALENT OR VIRTUAL HEIGHT IN KM
1200
1000
800
600
400
200
0
BASE-LINE

(A) SKETCH OF OSCILLOGRAM SHOWING TRANSMITTED PULSE AND SUCCESSION OF ECHOES

(B) UPPER PORTION OF ECHO-DEFLECTION CONDENSED BY LENS TO SINGLE LINE

(C) SUCCESSION OF FOCUSED ECHO-DEFLECTIONS ON MOVING FILM

(D) MULTIFREQUENCY RECORD OBTAINED WHEN WAVE-FREQUENCY IS INCREASED IN INCREMENTS BETWEEN SUCCESSIVE PULSE-TRANSMISSIONS

FIG. 1. Diagrammatic sketch showing formation of multifrequency ionospheric records.

Huancayo Magnetic Observatory these special conditions exist. Fig. 3 shows the result of experiments conducted by my colleague, Mr. Wells. An antenna placed in the magnetic north-south plane receives only one component; in the east-west plane only the other is observed, while at any intermediate angle both components are evident. Thus the performance required by theory is fulfilled. In these and other experiments the reasoning of the theory has been tested step by step until little doubt remains of its essential validity. One important question is yet to be tested, and the result will not affect our computations too seriously. It relates to the value in the outer atmosphere of a physical constant which we may call the "Lorentz polarization-correction." Some experiments on this point have been made, but again the theory tells us that we may sometime settle the question conclusively by making measurements on some future expedition which will carry us very close to the north or south magnetic poles.

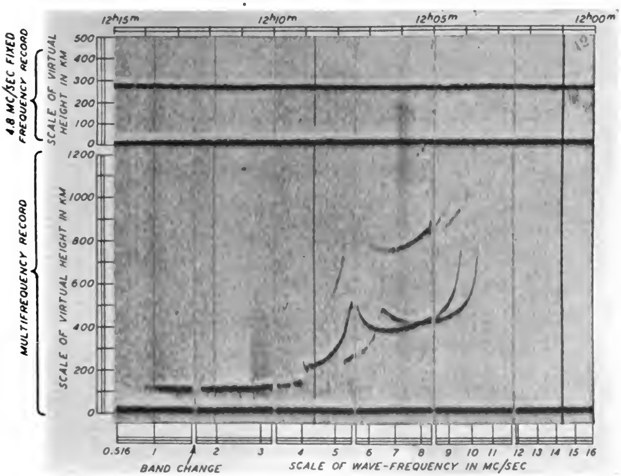


Fig. 2. Typical ionospheric record near noon in summer at midlatitudes, Watheroo Magnetic Observatory, November 26, 1939.

I mention proof of the theory in some detail for it permits exact quantitative measurements in atmospheric regions far removed

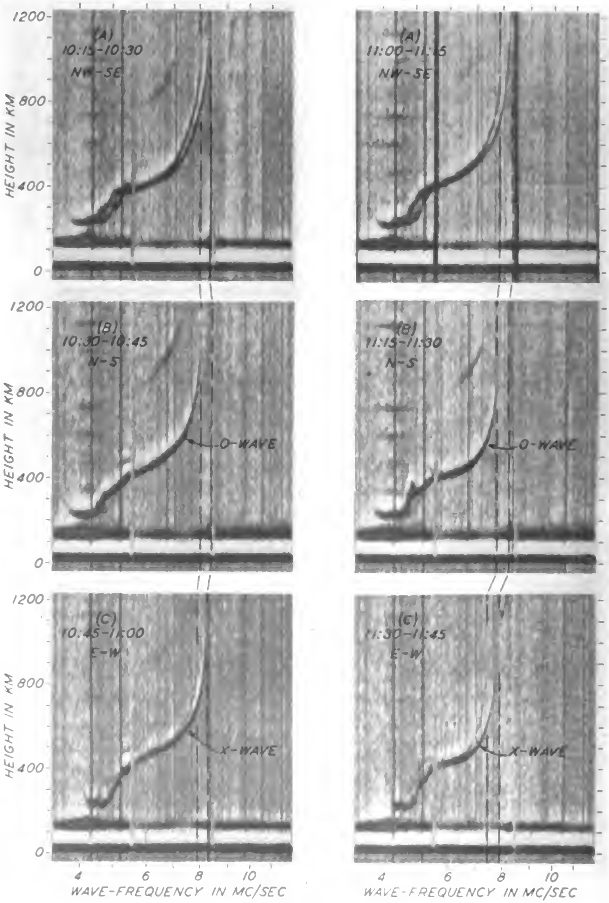


FIG. 3

Fig. 3. Polarization tests at geomagnetic equator with rotatable antenna, Huancayo Magnetic Observatory, January 11, 1940 (antenna orientation with respect to Earth's magnetic field: *N-S*, north-south; *E-W*, east-west; *NW-SE*, northwest-southeast).

from the Earth. In this proof the geomagnetic field has played the essential rôle. But it now seems probable that suitable experiments with improved apparatus will reverse this rôle and that variations of the geomagnetic field may be measured with precision at great heights by ionospheric observation. For example, suppose the magnetic diurnal variation be measured at a certain height above the Earth's surface. If at this height we are above the level of the electric currents producing this change, the diurnal variation will oppose that observed at the Earth's surface. In this way we may reasonably expect to determine the level of flow of the great electric current-systems in our outer atmosphere.

There are three important ionized layers in the outer atmosphere under direct influence of the Sun. These are the E-layer with maximum density of electrons at a level of about 100 km, the F_1-layer at about 225 km, and the F_2-layer at some 350 km. The two upper layers do not exist separately everywhere, but merge where the Sun is low and form a single layer in the night. This is illustrated in Fig. 4, where the F_2 layer appears as an enormous bulge of ionization above the F_1-layer, reaching a maximum of height directly under the Sun. Clearly the heights of the layers are not constant but change with time of day and season,

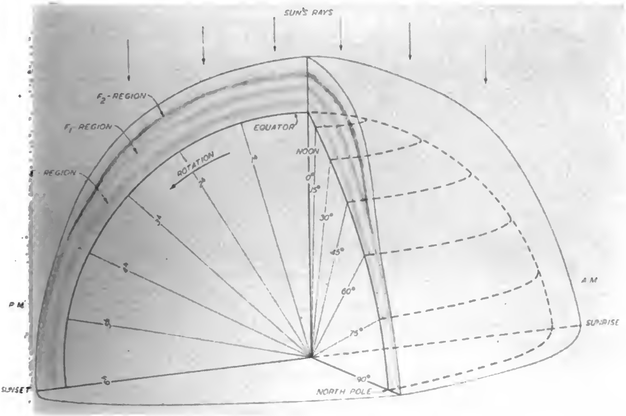


FIG. 4. Structure of ionosphere in sunlit hemisphere.

with latitude, and even with time along the sunspot-cycle. The density of ionization of the layers similarly changes, for predominantly the free electrons of the layers are produced by interaction of the Sun's ultraviolet light with the molecules along the light-path. At night the ion-density diminishes.

Dependence of outer atmospheric ion-density upon sunspot-activity is very marked. This is illustrated in Fig. 5. Here the average annual maximum density of electrons in F_2-region at noon as measured at Huancayo is compared with annual averages of sunspot-numbers. That the two agree in form of variation can hardly be doubted. From sunspot-minimum to maximum the change of average annual ion-density was about one to three for F_2-region. The lower regions show a change, similar in form, but of smaller magnitude. Over the same interval the E- and F_1-regions increased in density also, but only by some 50 per cent. This distinction is important to interpretation of geomagnetic effects.

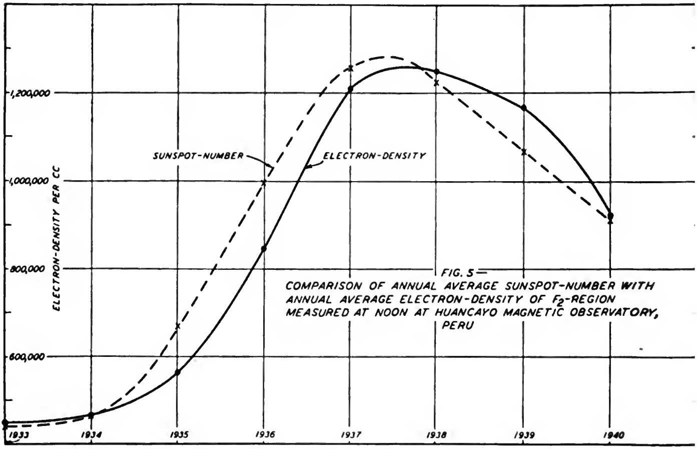


Fig. 5. Relation between electron-density of F_2-region and sunspot-numbers, 1933–40.

What evidence relating to geomagnetic diurnal variation can be obtained from the ionosphere? The most obvious is, of course, the distribution of ions and electrons in the various layers. Only in regions where sufficient electrification exists can there flow electrical currents to produce the geomagnetic change. But ion-density alone does not entirely determine the matter. The contribution of each ion to the electrical conductivity must be considered, and this depends on temperature and pressure of the atmosphere, on the masses associated with the changes, and on the strength of the geomagnetic field. All these taken together indicate that most suitable conditions for maximum current-flow probably exist at a level around 70 or 80 km above the Earth's surface. Experiments have been suggested, but not yet made, whose results should lead to quite precise estimates of electrical conductivity at the several heights.

The geomagnetic diurnal-variation changes in value by some 50 per cent from minimum to maximum of sunspot-activity. This corresponds under like conditions to observed changes in ion-density of E- and F_1-regions, but is greatly different from the change of some 300 per cent in F_2-region ion-density. This evidence would seem to exclude the high F_2-region as an important factor in the geomagnetic diurnal variation.

Evidence during radio fade-outs, first pointed out by Dellinger, is of almost conclusive importance (Fig. 6). The radio fade-out is coincident with a bright eruption in the solar chromosphere, and with a geomagnetic change of unique character (for example, see Fig. 15 of paper by J. A. Fleming, page 291). Ultraviolet light emanating from the eruption produces intense ionization of the lower E-region where the radio waves are absorbed, but the upper ionospheric layers are left relatively undisturbed. McNish has shown that the unique geomagnetic change associated with the fade-out is an augmentation of diurnal variation (Fig. 7). That is, during the fade-out, geomagnetic diurnal variation increases because of increased conductivity of the current-path. Because the atmospheric region below 100 km appears almost solely affected during the fade-out, it seems probable that not only the unique geomagnetic pulse but also the whole diurnal variation, of which the pulse is but an augmentation, arises from electric current-flow at about these levels.

Ionospheric effects associated with geomagnetic storms appear predominantly in the F-regions. Detailed investigations of these interesting phenomena are of recent date, since the introduction of modern automatic ionospheric recording methods. During magnetic disturbance, the ionospheric effects are so varied as to defy complete description here. There are, nevertheless, certain definite singularities during severe disturbances which can be illustrated by reference to the great magnetic storm of March 24, 1940. First we see from Figs. 8–10 that the ionosphere felt the impact of the storm at its commencement simultaneously on both sides of the Earth. At the Huancayo and Watheroo magnetic observatories, in almost opposite longitudes, the time of commencement was the same as nearly as could be determined. During the development of the storm the F-regions simply blew up and out. At Huancayo (midday) the F_2-region disappeared at great heights to be eventually replaced by a new F_2-region under the action of the Sun. At Watheroo, the F-region broke up into small ion-clouds within the hour. From this it must not be inferred that changes associated with magnetic activity are confined to the F-

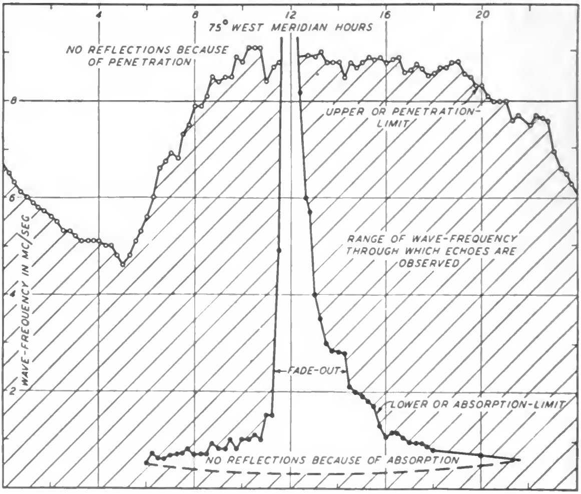


FIG. 6. Wave-frequency boundaries given reflection-limitation during fade-out; illustrating upward projection of lower or absorption-limit for fade-out of July 31, 1937, Kensington, Maryland.

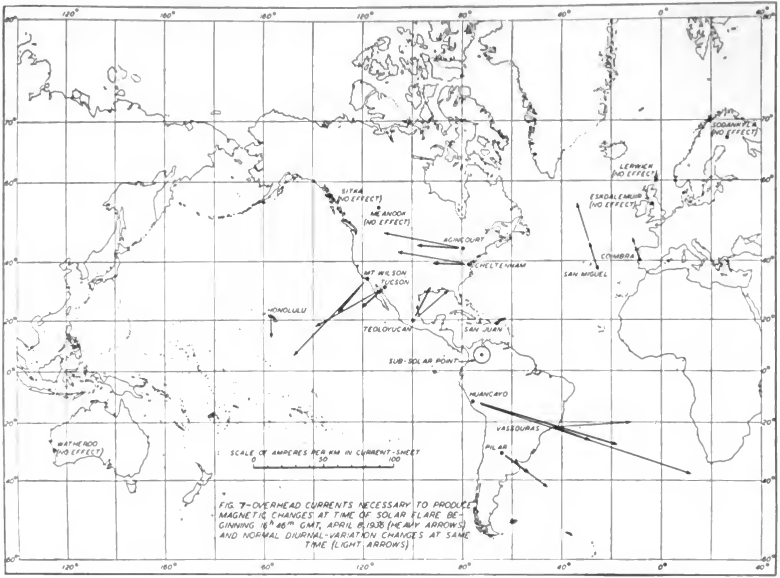


Fig. 7. Geomagnetic vectors during fade-out compared with normal diurnal-variation changes.

regions for increased absorption is evident at the 100-km level. But the most spectacular effects are observed at the higher levels. The effect of such disturbances on radio transmission can be imagined, for the radio waves must be scattered back and forth between small ion-clouds until much of the energy is lost.

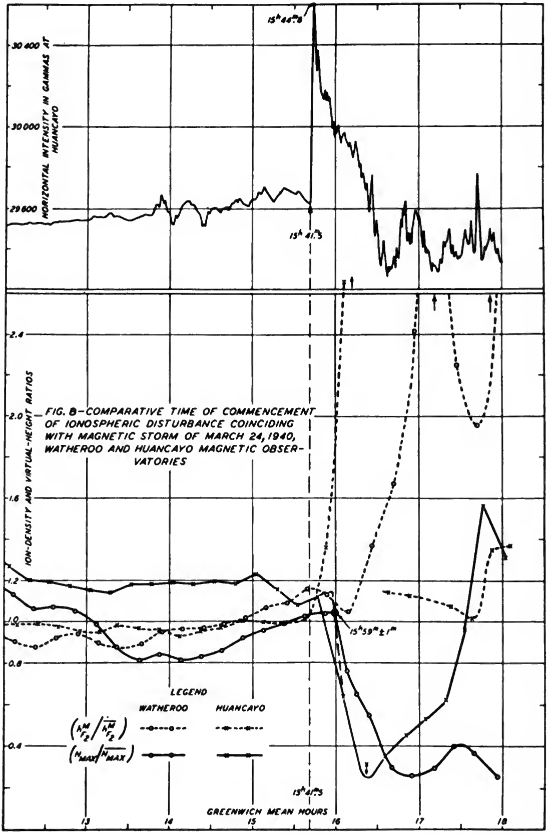


Fig. 8. Simultaneity of ionospheric disturbance with geomagnetic storm.

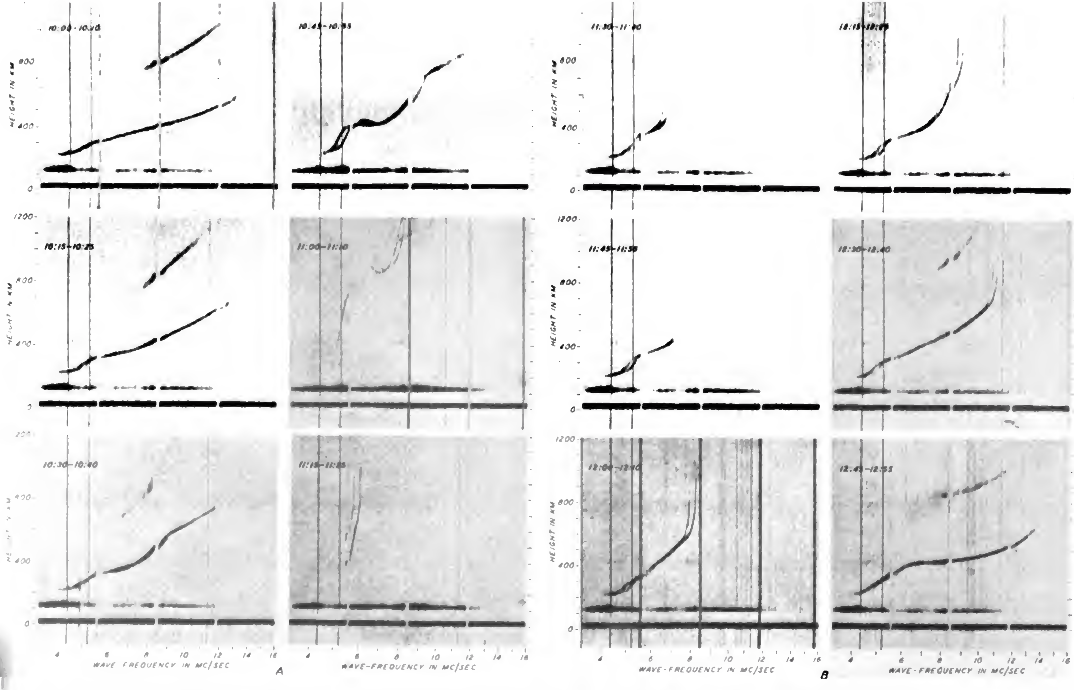


FIG. 9. Ionospheric records illustrating changes during geomagnetic storm of March 24, 1940, Huancayo Magnetic Observatory: *A*, disappearance of F_2-region; *B*, production of new F_2-region.

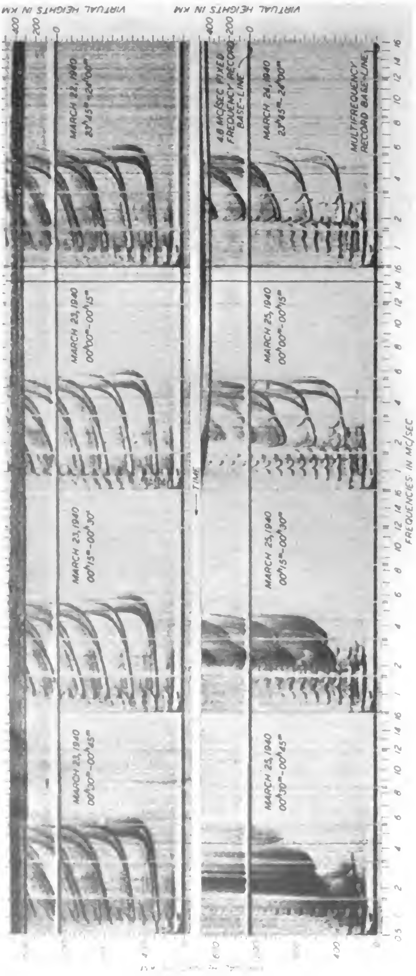


FIG. 10. Records illustrating commencement of ionospheric disturbance at Watheroo Magnetic Observatory coincident with geomagnetic storm of March 24–25, 1940, compared with quiet conditions of March 22–23, 1940.

The nature of the relation between ionospheric and magnetic disturbances is not yet evident. It may be that ionospheric changes arise in part from potentials induced by the magnetic field-change. On the other hand, it seems probable that some changes are produced by direct bombardment of the outer atmosphere by corpuscular radiation emanating from the Sun. Finally it may be that the ionospheric changes are effective in influencing the geomagnetic fluctuations. Whether the ring current-system associated with the storm is in the ionosphere, or above it, is not yet known. It may be said with assurance, however, that this part of the investigation of geomagnetism has only begun.

CORRELATIONS OF SHORT WAVE RADIO TRANSMISSION ACROSS THE ATLANTIC WITH MAGNETIC CONDITIONS

H. E. HALLBORG

Research Engineer, R.C.A. Communications, Inc.

(Read February 15, 1941, in Symposium on Geomagnetism)

I. Introduction

Short wave radio transmission, and the concurrent magnetic fluctuations in similar latitudes are intimately associated with changing rotational radiations from the sun. In fact, three broad fields of research, namely, the ionosphere, short wave radio transmission and geomagnetism have characteristics and properties due to a common origin, the sun.

It is well known that short wave radio transmission is largely controlled by the diurnal and seasonal characteristics of the ionosphere. It will be the purpose of this paper, however, to illustrate some of the most outstanding correlations, which have been observed by the author, connecting long distance short wave radio transmission with geomagnetism. The major interest will be centered in the Atlantic, in high geomagnetic latitudes; but since a broader perspective of short wave radio transmission is desirable a comparison of conditions at both high and low geomagnetic latitudes will be included. For this reason, short wave transmissions to New York from North Europe, South America and San Francisco, and from the Orient to San Francisco are included.

The apparent relationship which has been observed to obtain between magnetic disturbance, signal and power over the North Atlantic, and at lower geomagnetic latitudes, will be briefly reviewed. The latitude relationship is deduced from H intensity range data for all North American magnetic observatories made available to R. C. A. C. through the courtesy of the U. S. Coast and Geodetic Survey.

II. Measurement Methods

Simple and readily applied units of measurement are a requisite to world wide correlation. It was recognized that the most prolific source for short wave transmission data would be from normal operation on commercial world wide traffic. This requirement led to the application of a number rating system allowing the receiving technicians to estimate transmission conditions, on each eight hour watch, on selected circuits. The disturbance number rating system, adopted by R. C. A. C., and used since 1938, is shown in Table I.

TABLE I

Circuit Rating Number	Disturbance Condition	Signal Condition
0	Unusually quiet	Unusual strength
1	Normal	Strength and fading normal
2	Slight	Slightly below normal
3	Moderate	Considerably below normal
4	Severe	Nearly out; but still audible
5	Complete drop-out.	Inaudible

Earth current variations at the R. C. A. C. Riverhead Recording Station are measured in hourly units of "percentage variability." This unit is an expression for the measured length of the earth current trace per hour in terms of hourly baseline. For instance, "100 per cent variability" means that the measured earth current trace length is twice the baseline length. Percentage variability is an expression which is proportional to the mean rate of change of the earth field during the hour.

Cheltenham magnetograms are analyzed for hourly departures of the H intensity range by the customary procedure. A running chart is issued covering each solar rotation of 27 days. This chart divides each day into 4 six hour periods. The sum of the H ranges in each 6 hour period is plotted as ordinates in 27 day sequences.

The Cheltenham magnetic activity chart provides for a very considerable degree of circuit condition prediction since many solar disturbances persist for 6 or more rotations of the sun. The chart is a continuous record of solar, magnetic and ionosphere activity. The form of the chart is shown in the upper portion of Fig. 10.

A reproduction to scale of Riverhead Earth Current traces corresponding to R. C. A. C. Circuit Disturbance Rating Numbers 0 to 5 is shown in Fig. 1. Under each rating number is written the observed mean percentage variability corresponding to that number. The sketch above each number illustrates one equivalent shape of the trace.

It is of interest to compare the new 3 Hour K index method of magnetic rating, with the equivalent R. C. A. C. circuit disturbance ratings. This comparison is made in Fig. 2. The ordinates in this Figure are Cheltenham H intensity ranges, and the abscissas R. C. A. C. disturbance and K index ratings. A considerable band of overlap is seen to exist. The practical result of this overlap between mutual values 0 to 5 is that R. C. A. C. disturbance ratings and K values are interchangeable. All values of K above 5 will still produce R. C. A. C. rating 5—namely, complete signal drop-out or inaudibility.

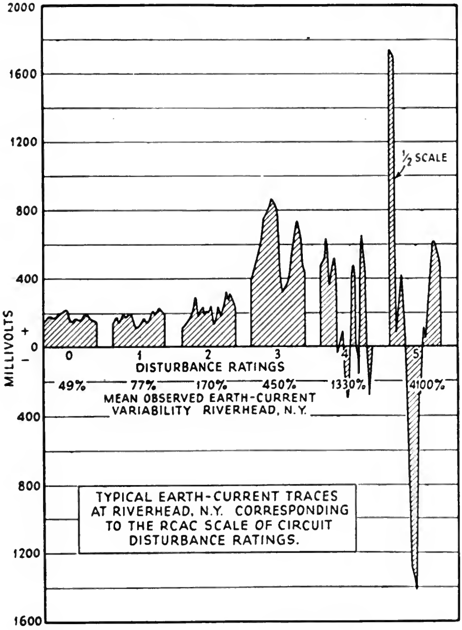


FIG. 1.

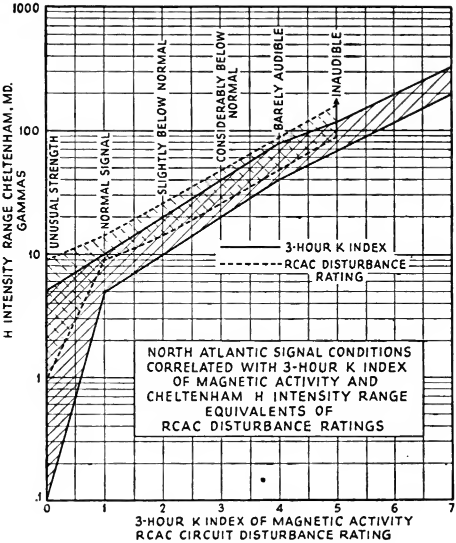


FIG. 2.

III. EARTH CURRENT VARIABILITY AND SIGNAL CORRELATION

London Signal "GLH," 13525 Kc, was continuously recorded during the month of April, 1936 at Riverhead, N. Y., and compared with the corresponding Riverhead earth current trace variability. Both quantities were reduced to a mean daily figure. The result of signal vs variability observations is shown in Fig. 3. The days of the month are identified by numbers, and the magnetic condition of each day by the magnetic legend. Three broad types of days were segregated, namely, quiet, storm and post-storm days. The full line curve is the form the points would have assumed if they follow precisely the inverse variability law. If post-storm days are neglected, during which period the signal is controlled by absorption, the inverse variability law is a fair mean of all observations. The decibels above and below normal signal in terms of disturbance ratings are indicated in the lower portion of Fig. 3.

This signal inverse variability relation has provided a means for computing the range of power that will permit commercial signaling across the Atlantic for various degrees of magnetic disturbance. This subject is considered in more detail in the next Section.

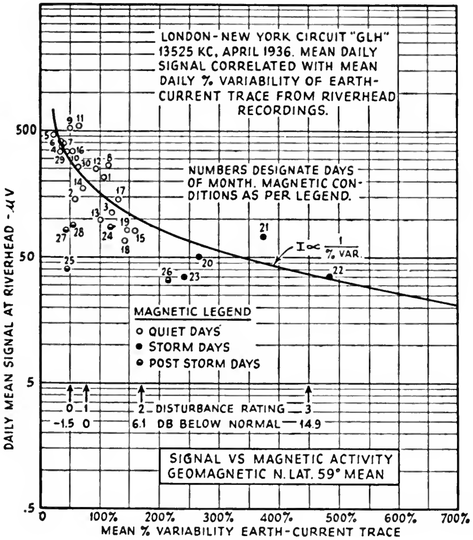


Fig. 3.

IV. Antenna Power vs Magnetic Disturbance at Geomagnetic Latitude 59° N

It was indicated in Fig. 3 that the order of signal level below normal for circuit rating 3 was 14.9 db. Experience has indicated that during such disturbance level over the North Atlantic the standard 20/40 KW R. C. A. transmitter will be operating at maximum rated power, namely, 40 KW, to avoid circuit interruption. If, therefore, 40 KW, and circuit rating 3, be taken as unity, and if the power varies inversely as the earth current variability, the power range for all disturbance ratings will vary in accordance with Fig. 4.

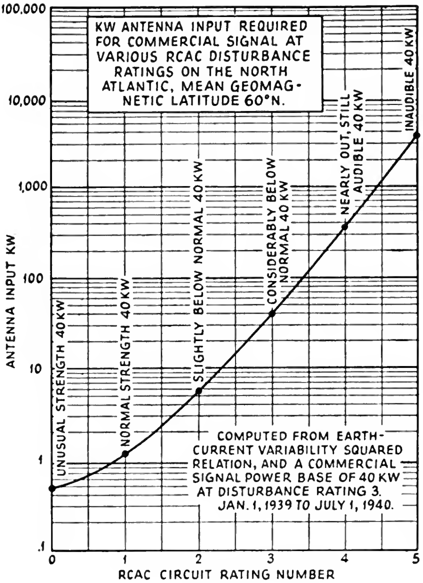


Fig. 4.

The extreme range of power will consequently be—

Rating	KW	Power Range
5	3460	7200
0	.48	

V. The Significance of Geomagnetic Latitude in Short Wave Transmission

The geomagnetic polar region is a zone of high magnetic activity, consequently of severe short wave fading. The equatorial region, on the other hand, is a belt of low magnetic activity, consequently of low short wave fading.

A short wave radio circuit will respond to the above geophysical relation in fair agreement with its mean geomagnetic latitude, particularly during magnetic disturbances.

In Fig. 5 is shown the mean daily circuit ratings observed during 16 major magnetic storms of 1939 and 1940, as well as the conditions 6 days before and after these storms on circuits having mean geomagnetic latitudes 14° N, 48° N and 59° N. The comparative disturbances are proportional to the total indicated shaded areas at the different latitudes. This area is very much lower at 14° N than at 59° N geomagnetic latitude.

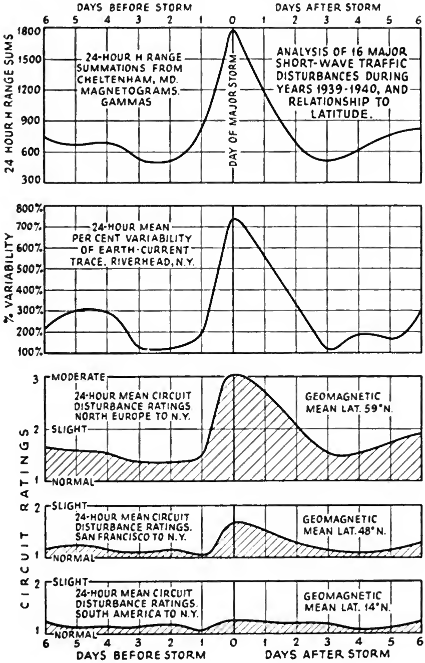


Fig. 5.

VI. Transmitting Power vs Geomagnetic Latitude

Cheltenham H intensity ranges correlated with Riverhead earth current variability has indicated that transmitting power may be related to H intensity range by the condition that the power required at any mean latitude varies approximately as the 3.5 power of the H range at that latitude. If this relationship is applied to the H range data of North American magnetic observatories of the U. S. Coast and Geodetic Survey for the year 1939 there is obtained the transmitting power vs geomagnetic latitude graph shown in Fig. 6.

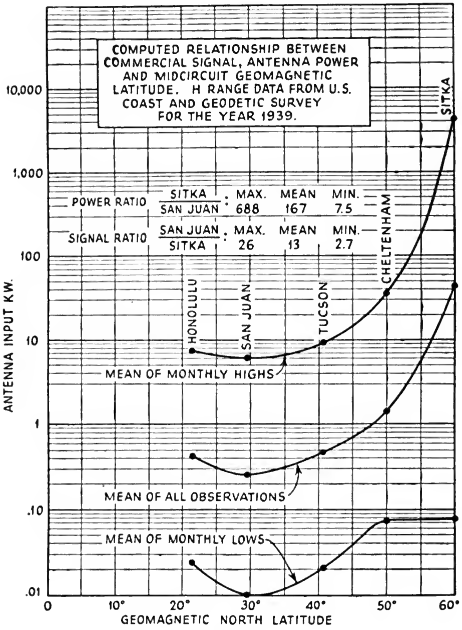


Fig. 6.

Comparison of the power requirements at geomagnetic latitudes 30° and 60°, from Fig. 6, produces the result shown in Table II.

It is of interest to compare these transmitting power computations vs geomagnetic latitude, from magnetic data, with the results of the long series of international broadcast recordings, at broad-

TABLE II

Geomagnetic Latitude	Antenna KW for Commercial Transmission		
	Max.	Min.	Yearly Mean (1939)
30°	6.15 KW	.010 KW	.265 KW
60°	4224 KW	.075 KW	44.3
Ratio 60°/30°	688	7.5	167

cast frequencies, published by J. H. Dellinger and A. T. Cosentino in the *Proceedings of the I. R. E.*, October, 1940.

Dellinger and Cosentino found that broadcast intensities, between North and South America, averaged 25 times those between North America and Europe, and were only 1/15 as variable.

The signal ratios at short waves, from Fig. 6, will be proportional to the square roots of the power ratios. This produces the relationship for the year 1939 shown in Table III.

TABLE III

Geomagnetic Latitude Ratio	Signal Ratio for Equal Power		
	Max.	Min.	Yearly Mean
30°/60°	26	2.7	13

These figures indicate that geomagnetic range is a major factor in international short wave transmission. A similar conclusion was reached by Dellinger and Cosentino for international transmission at broadcast frequencies.

VII. SHORT WAVE CIRCUIT INTERRUPTIONS VS GEOMAGNETIC LATITUDE

Circuit interruptions are definitely a function of geomagnetic latitude. Comparisons of the relative lengths of short wave circuit interruptions, on the particular circuits used in this analysis, during and after 16 major magnetic storms of 1939 and 1940 have produced the results shown in Fig. 7. A pronounced "knee" in the interruptions vs geomagnetic latitude graphs is seen to occur at 50° N geomagnetic latitude. The New York-San Francisco circuit (mean geomagnetic latitude 48° N) is less disturbed during the peak of the storm than is the New York-London circuit (mean geomagnetic latitude 59° N) 3 days following. The New York-Buenos Aires circuit (mean geomagnetic latitude 14° N) is comparatively undisturbed. There is some evidence of a slow southward disturbance filtration reaching a maximum 3 days after the storm.

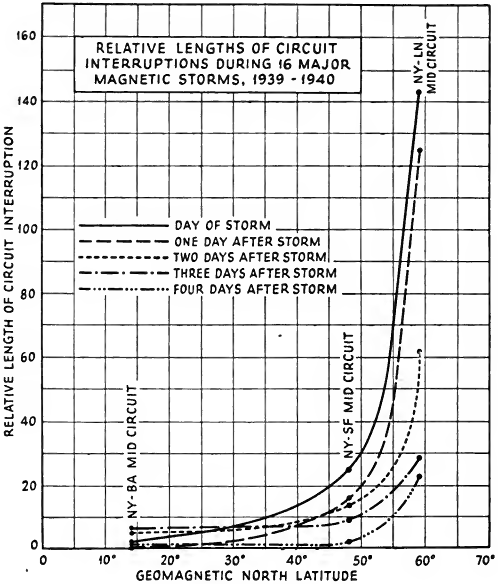


Fig. 7.

It is not to be assumed that all short wave transmission across the North Atlantic encounters the 50° N geomagnetic latitude "knee." The records show this was true a negligible percentage of total circuit time during the peak magnetic year of 1939.

VIII. Short Wave Transmission Conditions During the Solar Rotation February 6th to March 4th, 1939

A typical R. C. A. C. circuit disturbance rating graph is shown in Fig. 8. It covers the solar rotational period of February 6 to

March 4th, 1939. This period contained a wide range of magnetic disturbances over the North Atlantic circuits.

The plot at the top of Fig. 8 is the 8 hour average percentage variability of the recorded earth current trace at Riverhead,

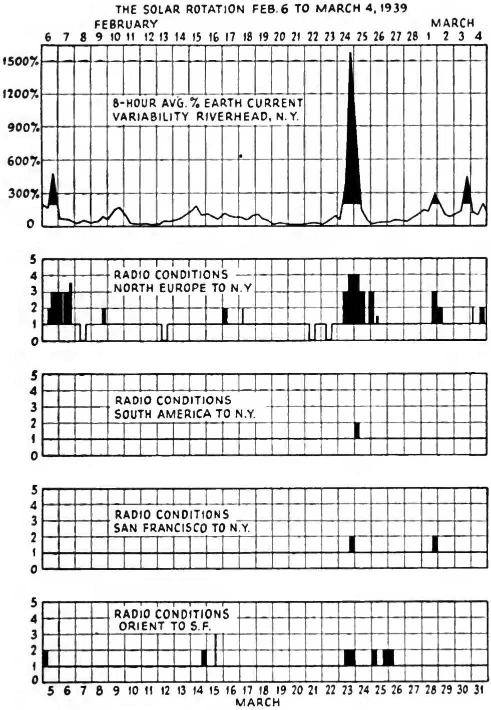


Fig. 8.

N. Y. Solid fills above 200 per cent indicate watches, or portions thereof, above the circuit disturbance level.

The 4 plots below show the reported scales of radio disturbance, in order, on the following circuits—

a. North Europe to New York.
b. South America to New York.
c. San Francisco to New York.
d. Orient to San Francisco.

Short period disturbances are indicated by thin lines of width proportional to their durations. A solid fill for each watch indicates an average disturbance for the watch of the amplitude indicated by the rating number.

The dates at the bottom of the chart are those for the oncoming solar rotation. If the disturbance areas on the sun persist, they represent dates of probable recurrence.

The North Europe disturbances during this rotation contain two characteristics of particular interest—

1. On February 22nd and 23rd, two days of unusual signal strength preceding the major storm day of February 24th.
2. On February 7th and February 26th, a persistence of circuit disturbance after earth current activity had dropped to normal.

The latitude effect, previously illustrated in Fig. 7, is quite in evidence in Fig. 8. The tranquility of the South America to New York circuit is outstanding.

IX. SHORT WAVE TRANSMISSION CONDITIONS DURING THE SOLAR ROTATION MARCH 17TH TO APRIL 12TH, 1940

Figure 9 illustrates perhaps one of the most exciting solar rotations in short wave history. It covers the period from March 17th to April 12th, 1940. This solar rotation created aurora, abnormal power and radio transmission disturbances, and wide comment in the public press.

The range of the percentage variability of the Riverhead earth current trace at the top of Fig. 9 has been extended to permit a visual estimate of its amplitude, above normal circuit disturbance level, particularly on March 24th and March 29th.

Features of particular interest were as follows—

a. A premonitory round the world short period fade, 1800–1900 GMT, on March 19th.
b. During the maxima of the disturbance peaks, March 24th and 29th, short wave communication over the North

Atlantic to Europe was completely cut off. European service was however continued without interruption by the relaying of messages to Europe via Buenos Aires.

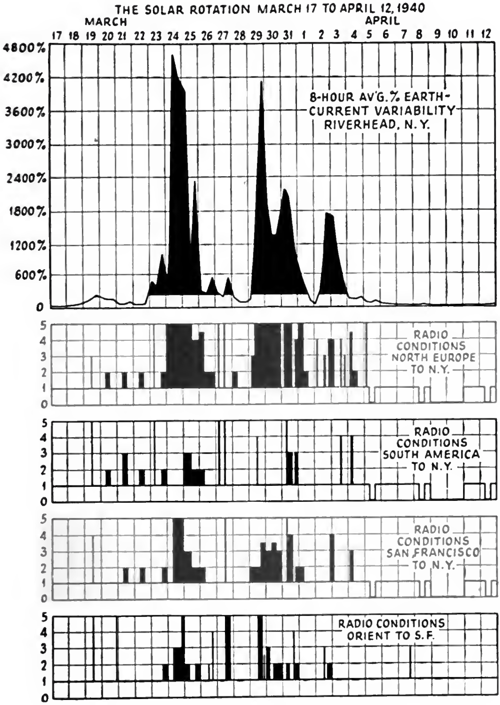


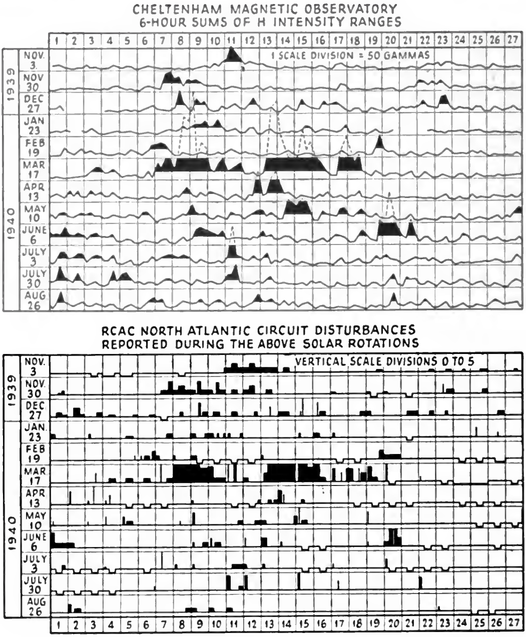
FIG. 9.

c. A second round the world fade, 1615–2000 GMT, on March 27th, preceded the second magnetic storm on March 29th.

d. The South American circuit again has the helpful influence of low geomagnetic latitude. It was entirely normal during the storm peaks.

X. North Atlantic Transmission and Cheltenham H Ranges Compared over a Period of Twelve Solar Rotations

North Atlantic short wave circuit disturbances and Cheltenham H intensity ranges are plotted in 27 day solar sequences for a period of 12 solar rotations, beginning November 3, 1939, and ending September 21, 1940, in Fig. 10. The date of commencement of each solar rotation is at the left of each row, and corresponds with horizontal numeral number 1.


CHELTENHAM MAGNETIC OBSERVATORY
6-HOUR SUMS OF H INTENSITY RANGES

RCAC NORTH ATLANTIC CIRCUIT DISTURBANCES
REPORTED DURING THE ABOVE SOLAR ROTATIONS

Fig. 10.

Cheltenham H ranges have been corrected for scale calibrations, and each point is a six hour sum of hourly departures. The vertical scale is 50 gammas per division. Values above 250 gammas are indicated by dotted lines.

The North Atlantic circuit disturbances are represented by disturbance numbers 0 to 5 in accordance with Table I. The bottom line of each row represents 0 rating, the top line 5 rating.

The charted time interval selected includes the peak of the magnetic disturbance in the current eleven year cycle. This chart is useful for the purpose of indicating concurrent radio and geomagnetic disturbance; but may produce a misleading impression as to the value of such charts for disturbance prediction. The selection of a period of twelve solar rotations earlier would have shown repetitive disturbances for as many as ten solar rotations.

A careful study of these parallel charts in Fig. 10 will reveal the persistence of radio disturbance, the repetitive property of moderate storms and the non-repetitive property of major storms. It will bring to light one anomaly, the first to come to the author's attention, namely, the association of excellent signal conditions, in the face of considerable magnetic disturbance during the first four or five days of the sequence beginning July 30, 1940.

If the question of location of a monitoring magnetic observatory for any long distance short wave circuit is considered it is now evident that the ideal location for such an observatory would be at the mid-point of the circuit. Such a site selection presents practical difficulties when the application is in the mid-Atlantic. It will also be evident that the selection of any site, on land, above 60° geomagnetic latitude will provide much information of scientific value, both as to geomagnetism, and the ionosphere.

XI. Discussion and Conclusion

The data here presented is further evidence that short wave radio transmission, the ionosphere and geomagnetism are closely related to the changing rotational complexion of the solar surface.

The closer scientific correlation of geomagnetism and radio transmission, and the ionosphere, at locations more suited to bring out the latitude characteristics, will further supplement our knowledge of all these fields.

There is shown to exist during geomagnetic disturbances a short wave communications "knee" at about 50° geomagnetic North latitude. Great circle paths intercepting above 50° do so at the cost of increased power and circuit interruption.

On the other hand, great circle paths that follow the lower latitudes are relatively free from circuit interruption during geomagnetic disturbances and may operate at lower power with lesser fading, in proportion to the long observed improvement in geomagnetic variability at the lower latitudes.

MAGNETISM AND ITS USES

PAUL R. HEYL

Physicist, National Bureau of Standards

(Read at Girard College February 15, 1941, in Commemoration of A. D. Bache)

It has been my experience, when confronted with an intricate and highly developed scientific subject, that the best line to follow is the historical approach. However complicated such a subject may appear in its present state of development, its beginnings were much simpler. A few outstanding facts attract the attention of the first observers, and a comparatively simple theory is devised to coordinate and account for these phenomena. Further study on the part of the curious brings to light additional facts, less obvious than those first observed, and theory has to be modified to include these new discoveries. Again and again this process is repeated until the subject embodies such a formidable mass of interconnected detail that it is difficult to get it in focus and perspective without retracing step by step its historical development; and along this line we shall approach the subject of Magnetism and its Uses.

The first magnet known to man was the loadstone, a naturally magnetic ore of iron which occurs in many places all over the earth, among them a district in Asia Minor called in ancient times "Magnesia," whence, by our heritage of Greek tradition, we have our word "magnet." This stone differs from all other stones in that it possesses the power of attracting to itself pieces of iron. It is therefore almost certain that early man could not have recognized the peculiarity of this stone until he had metallic iron at his disposal. I say "almost" because there is a bare possibility that he might have noticed that the stone would attract to itself broken bits of its own substance. There is, however, no mention of such a phenomenon by any writer earlier than the thirteenth century A.D. We may therefore safely date the beginnings of magnetic science not earlier than the Iron Age.

The transition from the Bronze Age to the Age of Iron occurred, of course, at different times in different places. In the Mediterranean basin, where our records are most complete and trustworthy, iron was apparently well known before 1000 B.C. The Homeric poems contain references to armor and weapons of both brass and iron, showing that the newer metal had not yet completely displaced the older. There is evidence that the iron there mentioned was at least sometimes a crude variety of steel. In the ninth book of the Odyssey the poet tells how Ulysses and his companions blinded the one-eyed giant Polyphemus by thrusting into his eye a burning stake. The passage reads (in Butcher and Lang's prose translation):

> And as when a smith dips an axe or adze in chill water with a great hissing, when he would temper it—for hereby anon comes the strength of iron—even so did his eye hiss round the stake of olive.

Once man had iron at his disposal it could not have been long before he noticed the attractive power of the loadstone. This discovery was probably made independently in different localities. The Greeks had a legend that a certain shepherd noticed that the iron tip of his staff adhered to some rocks over which he was walking. In this or in other ways that can equally easily be imagined, this important discovery might readily have been made.

When a new phenomenon is observed the first question that comes into the mind is "Why?" To account for the attractive power of the loadstone there was early formulated a simple theory. This was that the loadstone was inhabited by an indwelling soul or spirit which enabled it to move the iron. That this spirit dwelt in the loadstone rather than in the attracted iron was obvious to these early observers because iron itself would not attract other pieces of iron. To this one exception was early noted.

In the island of Samothrace in the Ægean Sea, at some time prior to 400 B.C., there was invented a magnetic toy consisting of a loadstone and a set of small iron rings. If one of these rings was picked up by the loadstone this ring would at once acquire sufficient magnetism to support a second ring hanging beneath it. This second ring would likewise support a third ring, and so on, until a chain of perhaps four or five rings could be supported by the loadstone. If now the upper ring of the chain were detached from the loadstone the rings would lose their magnetism and the chain would fall apart. In modern parlance, this toy illustrates the difference between the permanent magnetism of the loadstone and the temporary, or induced magnetism of the iron rings.

Plato, the pupil of Socrates, relates how the Master used the Samothracian rings to point a moral to his disciples. As Socrates died in 399 B.C., we are able to set an approximate date for the invention of this magnetic toy. Plato's account is as follows:

> "The gift which you have of speaking excellently about Homer," says the sage, "is not an art, but, as I was just saying, an inspiration. There is a divinity moving you, like that in the stone which Euripedes calls a magnet, but which is commonly known as the stone of Heraclea. For that stone not only attracts iron rings, but also imparts to them similar power of attracting other rings; and sometimes you may see a number of pieces of iron and rings suspended from one another, so as to form quite a long chain; and all of these derive their powers of suspension from the original stone. Now, this is like the Muse who first gives to men inspiration herself, and from those inspired, her sons, a chain of other persons is suspended, who take inspiration from them."

The Samothracian rings produced a profound impression on the ancient world, as is evidenced by the number of post-Platonic writers who mention them. Eight hundred years after the death of Socrates the interest in this toy was unabated. About 400 A.D. some one showed these rings to St. Augustine in northern Africa, who, as he says himself, was thunderstruck by them. The good bishop goes on to say:

> Yet far more astonishing is what I heard about the stone from my brother in the episcopate, Severus, bishop of Milevis. He told me that Bathanarius, once Count of Africa, when the bishop was dining with him, produced a loadstone and held it under a silver plate on which he placed a bit of iron; then as he moved his hand, with the stone underneath the plate, the iron upon the plate moved about accordingly. The intervening silver was not affected at all, but precisely as the loadstone was moved backward and forward below it, no matter how quickly, so was the iron attracted above.

There is a still earlier reference to this phenomenon in the writings of the Roman poet Lucretius, who died about 55 B.C. He speaks of iron filings "raving within brass basins" if the stone is moved underneath. It was apparently an early discovery that the magnetic influence could be transmitted through non-magnetic bodies (silver and brass) without being apparently weakened.

With the fall of Rome and the onset of the Dark Ages in Europe, the interest in scientific experiment waned. The loadstone, however, maintained its hold on popular imagination. Many legends of its wonderful power arose, such as that of the magnetic mountain which could draw the nails from ships that ventured too close to it. Medieval physicians adopted the loadstone eagerly, and prescribed it for all sorts of disorders. It was applied in powdered form as an ingredient of a plaster to wounds caused by arrows, with the idea of drawing out the iron arrowhead from the body. A certain prince of India ordered that the dishes in which his food was prepared should be made of loadstone, in order to preserve his youthful vigor; and a grave doctor of those times administered doses of powdered loadstone to both parties in cases of estrangement between husbands and wives.

This latter procedure was not without justification, when seen against the background of the belief of the times. It was obvious that the patients in question were suffering from the lack of a certain spiritual principle which was believed to exist in the loadstone, and the administration of that medicine followed as naturally as a modern prescription of cod liver oil because of its vitamin content.

The belief in the medical virtues of the magnet persisted until recent times. "Magnetic" belts were worn to prevent disease in the nineteenth century, and within my own recollection there were those who believed that a bed should be set up in a north and south position in order that the sleeper might obtain the maximum benefit from the magnetic field of the earth.

The first 2000 years of the history of magnetism, which we have briefly reviewed, produced the discovery of several fundamental facts and principles, but brought about no practical application of this knowledge. Not until the twelfth century do we find the first useful application of magnetism—the mariner's compass.

Where and by whom the compass was invented is unknown. It has been ascribed to the Chinese, the Arabs, the Greeks, the Etruscans, the Finns and the Italians. There is no lack of legend on the subject, but modern historical scholarship regards such legends as mythical. The first unquestionable reference to the use of the compass in Europe is found in the writings of an English monk, Alexander Neckam, in the year 1200. Neckam speaks of an instrument consisting of a magnetized needle mounted on a pivot, which sailors use in cloudy weather, by day or night, when they lack the guidance of the sun or the North Star. From this time onward references to the compass are numerous, and in consequence it is believed that the invention of the compass occurred some time in the latter part of the twelfth century. It may well have happened that, the time being ripe, this invention was made independently in different places. This has been the case more than once in modern times.

Two varieties of the compass were in use in early times—the pivot type and the floating type. In the latter variety the needle was supported by a small float of wood or cork in a bowl of water. The needles at first in use could not be depended upon to retain their magnetism for long periods, as the art of steel making was as yet in an undeveloped state, and it was necessary for mariners to provide themselves with a loadstone to remagnetize the needle frequently.

The inevitable question, "Why does the needle point north?", received two answers. Some said that there existed in the far north a mountain of loadstone; others, probably influenced by the prevalent astrological ideas of the times, supposed the needle to be governed in some mysterious way by the pole star.

With the compass needle as an instrument of research, a number of new magnetic facts were discovered. Roger Bacon, the best known man of science in the thirteenth century, observed that the magnetism of the loadstone was concentrated in two places, which we now call the poles of the magnet. He noticed also that a needle which had been touched by the loadstone would afterward be attracted by the part of the stone which had touched it, but would be repelled by a certain other part. Other important discoveries were made by a Frenchman named Peter de Maricourt, called Peter Peregrinus, because he had made a pilgrimage to the Holy Land. His discoveries are described in a letter written by him in 1269.

In this letter Peregrinus describes the experiment of cutting a loadstone in two parts, with one of its poles in each part. Much to his surprise, he found that each part was now a complete magnet with two poles. On placing the two parts together again, the two new poles at the center disappeared, leaving the stone in the same magnetic condition as before it was cut. Moreover, he found that this process was capable of apparently endless repetition, any part of a magnet, however small, becoming a complete magnet in itself. Peregrinus found also that a magnetized needle, when brought near a loadstone without touching it, would assume a definite angular position with respect to the stone. Every student in the physical laboratory who maps a magnetic field by means of a compass needle, repeats this seven-century old experiment of Peregrinus.

The next important discovery in magnetism is generally, but not universally, attributed to Columbus, who, on his first voyage to America, found that the compass needle did not always and everywhere point to the exact north. This variation is now known by the name of the declination of the needle.

It appears that in the time of Columbus the needle in western Europe must have pointed so nearly to the true north that its declination was unsuspected. Small variations, if noticed, could readily have been explained away as due to imperfect construction, such as pivot friction. In addition, the thought of those days was profoundly influenced by astrology, and with this background there was an obvious reason why the needle ought to point north, and no reason at all why it should point anywhere else; for there in the north was the invariable pole star, while elsewhere in the heavens all the other stars rose and set, continually changing their positions and their supposed influence on mundane affairs.

The long westward journey of Columbus brought him eventually into a region where the declination of the needle was too great to be overlooked or explained away. His first observations were soon confirmed by other navigators, and the declination of the needle became a recognized scientific fact.

This discovery rendered untenable the hypothesis that the compass needle was governed by the pole star, but the other hypothesis of the magnetic mountain in the far north was rather strengthened by it; for all that was necessary was to suppose that this mountain was located at some distance from the geographical north pole. An observer in line with both mountain and pole would find his compass pointing to the true north, while another observer at some distance east or west of this line would notice that his compass pointed not to the pole, but to the mountain.

Scarcely had the nautical world become used to making the necessary allowance for the declination of the needle when it was again upset by the discovery that this declination was continually changing, at a slow rate, it is true, but quite perceptible over a period of years. For instance, it was found that in London, in 1580, the compass pointed as much as 11° east of north. Forty years later the declination was only 6°, and by 1657 the needle pointed exactly north. The declination then became westerly, reaching a maximum of 24° in 1820. The needle then began to move back toward the north, which it has not yet reached. If it does not change its mind in the meantime, it may again point north at London in another hundred years.

This discovery of the change in the declination rendered untenable, in its turn, the hypothesis that the compass needle pointed to a magnetic mountain—untenable, except by those possessed of the faith that could move mountains. Over two hundred years were destined to elapse before any satisfactory answer could be given to the question, "Why does the needle point north?"

Changes in declination take place all over the earth, and the importance to navigators of an exact knowledge of this variation is obvious. In recent times much attention has been given to the preparation of tables and charts of this important correction. For many years the Carnegie Institution of Washington maintained a special ship equipped with magnetic measuring instruments, which made a survey of the oceans of the world every ten years. A few years ago, while on one of its cruises, this ship was destroyed by an explosion. The British Admiralty then assumed this task, and built another ship for this purpose. Land stations for magnetic observations are also maintained by various nations.

Observations at these land stations have shown that the variation of the needle is itself rather irregular. Beside the long period variation already mentioned it has been found that there is an annual fluctuation of smaller amount, and also a variation from day to day. Moreover, this diurnal variation may be much greater on some days than on others, and it is now customary to speak of magnetically "quiet" days, and of "magnetic storms." The poetic phrase: "True as the needle to the pole" has lost its force for scientists.

As if all this were not enough, it has been found that there is still another peculiarity in the behavior of the compass needle. In 1544, Dr. George Hartmann, in Germany, in constructing a compass, balanced an unmagnetized needle carefully on its pivot, and then afterwards magnetized it. He was surprised to find that the magnetization had destroyed the equilibrium of the needle, as the north end now pointed downwards, and it was necessary to add a little wax to the south end to restore the equilibrium.

Some years later, Robert Norman, an English instrument maker, independently rediscovered this dip of the needle and investigated it carefully. He mounted a needle on a horizontal axis instead of the usual vertical pivot, to give the dip free play, and found that after magnetization the needle exhibited a dip of a little over 70° below the level line, thus taking a position more nearly vertical than horizontal. Continued observation showed that the dip of the needle, like its declination, differed at different places, and changed slowly with time.

We have now reviewed the history of magnetism from its discovery, some time prior to 1000 B.C., up to the beginning of the nineteenth century, a period of some three thousand years. We have seen how our knowledge of magnetism grew slowly, hundreds of years sometimes elapsing between the successive steps of progress, but that eventually quite a number of fundamental facts were established, and one practical application made—the compass. When we consider the meager equipment of these early observers it is indeed remarkable that they accomplished as much as they did. All that they had was a loadstone and a compass needle. The nineteenth century, however, was to witness a great stimulation and acceleration in the advance and application of our knowledge of magnetism, due to two discoveries of the first rank in importance.

Electricity has a history equally as old as that of magnetism, but until the nineteenth century the two had always been considered different in nature. In 1819 Hans Christian Oersted, a professor in the University at Copenhagen, discovered that an electric current could produce a magnetic effect, and twelve years later Faraday made the converse discovery that a magnet could produce an electric current. This union of these two subjects was remarkably fruitful. In addition to giving a great impetus to the development of theory, it made possible the whole of modern electrical engineering, with its manifold applications.

Oersted found that if a wire carrying an electric current were held parallel to a compass needle at a little distance above it, the needle would be deflected from the north, and that this deflection would be greater the stronger the current. This discovery, published in 1820, excited much interest in the scientific world, and five years later, Sturgeon, following this lead, wound a coil of wire around an iron bar, and on passing an electric current through the wire found that the iron became strongly magnetized, but lost its magnetism when the current stopped. Such an electromagnet could easily be made thousands of times as strong as any loadstone, and a new and powerful instrument of research was thus made available.

Faraday, influenced by the results of Oersted and Sturgeon, connected the ends of a coil of wire to a galvanometer, and placed within the coil a steel bar magnet. Contrary to his expectations, there was no permanent current while the magnet was at rest in the coil, but a temporary current was produced while it was being put in and taken out.

This experiment appears to have attracted attention in other than the scientific circles of the day, as is evidenced by some verse that has come down to us.

Around the magnet, Faraday
Is sure that Volta's lightnings play;
  To bring them out was his desire.
He took a lesson from the heart:
'Tis when we meet, 'tis when we part,
  Breaks forth the hid electric fire.

One result of experiments with the new powerful electromagnets was greatly to increase the number of bodies known to be magnetic. From ancient times only two such bodies had been known—iron and the loadstone, but when tested in strong magnetic fields nearly everything else was found to be feebly magnetic. A few substances, such as bismuth and gold, were found to be repelled instead of attracted by the magnet. Such bodies are now called "diamagnetic."

Many practical applications have been made of the fact that an electromagnet will lose and regain its magnetism as the exciting current stops and starts. When we think of what our modern world would be without the telegraph, the telephone, the electric bell, the electric motor or the dynamo, all of which contain an electromagnet as an essential part, we can realize the importance of Oersted's discovery. Denmark delights to honor her famous son, and Oersted Park is one of the beauty spots of modern Copenhagen.

But these modern applications of the electromagnet would have been of but limited use had it not been for Faraday's discovery. An electromagnet requires an electric current to operate it, and the only source of current available to Oersted or Faraday was the galvanic cell. When we think of the number of galvanic cells that would be required to operate the motor of, say, an electric elevator, we can see that on this basis the expense and trouble of large scale applications of Oersted's discovery would be prohibitive. It was Faraday's discovery that eventually made electric current cheap enough to make these applications practicable.

If we imagine Faraday moving his magnet in and out of the coil rapidly enough, the current impulses in the galvanometer would follow each other so closely as to be practically continuous. It is true that these successive impulses would flow in opposite directions, giving an alternating current, but if while working the magnet with one hand he reversed the galvanometer connections with the other, the alternating current would be rendered direct, or as we say nowadays, "rectified." This is the fundamental principle of the modern dynamo, by which mechanical power from the steam engine is made to furnish cheap electric current.

The first machines of this character followed Faraday in using a permanent steel magnet, but instead of moving the magnet in and out of a coil of wire the magnet was kept stationary, and the coil of wire rotated rapidly close to the magnet's pole. These were called "magneto-electric machines," or more briefly, "magnetos." The steady demand for machines of more current capacity required the use of stronger magnets, and the steel magnets were soon replaced by electromagnets. Such machines were called "dynamos," because of their greater power output.

The electromagnets of the first dynamos were excited by current from an outside source, such as a magneto, but it was soon found that a dynamo really needed no outside excitation. The iron of the electromagnet retained enough residual magnetism to generate a little current when started up again. By using the dynamo's own current to excite its magnet the output of the machine would rapidly build itself up to a maximum value, the limit being reached when the iron of the electromagnet became magnetically "saturated." The modern dynamo is thus a combination on a large scale of the discoveries of Oersted and of Faraday.

In the manufacture of steel, as practiced at the present time, application is made of magnetic principles in several interesting ways. In order to harden steel it must be heated to a certain definite temperature, and then quickly cooled. When a large piece of steel is being heated in a furnace the determination of this temperature is not a simple matter. Observations of the surface temperature can be made, but the determination of the temperature at the center of a large piece is a matter of some difficulty. However, all grades of steel lose their magnetism when heated sufficiently, and it so happens that, for certain grades of low-carbon steel, the critical temperature at which the magnetism vanishes is quite close to the hardening temperature. As long as there is any magnetism left, even in the inside of the steel, it can be detected by electrical methods, and apparatus has been invented by which magnetic tests can be made on a large piece of hot steel in a furnace.

Magnetic principles are used also in testing finished steel products. For instance, a railroad rail may contain internal cracks which no outside inspection can detect; and such cracks, subjected to the severe strain of daily use, may increase in size and become dangerous. But such cracks will alter the magnetic quality of the steel in their vicinity, and by passing the rail through a suitable magnetic testing instrument weak spots can be discovered. The same principle has been used for testing periodically the steel cables of hoisting machinery.

It is interesting to note in passing that steel, which owes something to magnetic science for the modern perfection of its manufacture, has shown itself in one respect rather ungrateful. The use of steel in the construction of ships interferes with the proper action of the magnetic compass, and makes necessary so many complicated and troublesome corrections that the magnetic compass is now being widely replaced by the non-magnetic gyrocompass. The time may not be far distant when this first practical application of magnetism may be as much of an antique as the spinning wheel.

In this brief review we have seen some of the many applications of magnetism which have been made possible by the discoveries of the nineteenth century. What, on the other hand, have these discoveries done to advance our theoretical knowledge of magnetism?

Much has been done on this side also. For one thing, the nineteenth century has given us a partial answer to the question, "Why does the compass needle point to the north?"

About the year 1600 the earth was believed to be a great permanent magnet like the loadstone. We have seen that the discovery of the variation of the compass needle made this hypothesis untenable, for there was then no reason known why the poles of a magnet should change their position. The discovery of the electromagnet in the nineteenth century suggested a partial solution for this difficulty, for it was found that the poles of an electromagnet would shift their position with a change in the position of the exciting coil.

For example, if we take a terrestrial globe made of iron, wind a coil of wire around it following the direction of its equator, and pass an electric current through the coil, the globe will become magnetized, and its magnetic poles will coincide with the geographical poles. But if the exciting coil be tilted slightly, the magnetic poles will suffer a corresponding change in position. Nor is it necessary that the coil be wound on an iron core. With a core of stone or wood, or even with no core at all, the magnetic field of the coil will still be in evidence, though with diminished intensity, and will follow any change in position of the coil. If, then, the earth were a huge electromagnet instead of a permanent magnet, the change of place of the magnetic poles might be possible. But where are we to look for the current to excite this great electromagnet?

For this question also, the nineteenth century provided a suggestion. When the first telegraph lines were established about the middle of the century it was found that there were certain days when the lines were unworkable. Electric currents of mysterious origin found their way into the lines and made it impossible to transmit messages. The origin of these currents was finally traced to the earth itself, the current entering the telegraph line through its grounded ends. Further study showed that these earth currents were always present, usually of an intensity too feeble to cause serious trouble; but on certain days the intensity of these currents would rise to such a point as to drown out the current carrying the message. It was further noticed that at such times there was apt to be an aurora visible at night, and that the variation of the compass was unusually erratic. Such disturbances are known by the name of "magnetic storms."

In addition to these always present earth currents, it is now recognized that there are electric currents in the upper atmosphere, which at times may become sufficiently intense to be accompanied by a glow discharge which we call an aurora, and the joint action of these earth and air currents of electricity is regarded as at least a considerable factor in the earth's magnetism.

But with all this progress in magnetic science, theoretical and applied, we are still without a satisfactory answer to the question, "What is magnetism?" The experiments of Oersted and Faraday opened our eyes to the close relation that exists between magnetism and electricity, but as we cannot yet answer the question, "What is electricity?", we have only reduced two mysteries to one. And yet to that extent we have simplified the task before us, and we continue our researches in the confident hope and belief that at any time some new discovery, as fruitful as those of Oersted and Faraday, may unlock for us this last mystery.

.t
a·

Copyrighted material

S. PIETRO IN VINCOLI AND THE TRIPARTITE TRANSEPT IN THE EARLY CHRISTIAN BASILICA[1]

RICHARD KRAUTHEIMER

Vassar College

(*Read April 1939*)

ABSTRACT

Although remodelled several times from the XV through the XVIII centuries, the church of S. Pietro in Vincoli in Rome preserves considerable remnants of its Early Christian structure. As can be shown from these remnants the edifice was originally a basilica with a nave bounded by two arcades of eleven arches each, with two aisles, a transept and one single apse. Traces of walls and of an arch above the XV century vaults indicate that in a first project the transept was to be divided into three parts: a center bay slightly wider than the nave was to communicate with two wings by a series of three arches on either side; these arcades were to carry upper walls with three openings each. While the building was still in progress this first project was replaced by an arrangement in which the three divisions of the transept were marked off only by four projecting piers. This was the arrangement preserved up to the XV century.

A comparative analysis of the documents and of the structure goes to prove that this basilica should be dated between 420 and 450. While the church was conecrated to the princes of the Apostles there are indications that it succeeded an older Christian building, possibly a community house dedicated to Saint Peter alone; from the VI century on the new church gradually changed its dedication back to Saint Peter. Remnants of older structures which were found in 1878 below the apse are published for the first time from a drawing in the Lanciani collection.

The tripartite transept is unique in Early Christian churches in Rome. A study of the transept of Early Christian times proves that three types have to be distinguished: the "continuous transept" which occurs only in some Roman basilicas of the IV century, the "cross transept" which is common in the Near Eastern Coastlands and the "tripartite transept" which is found in S. Pietro in Vincoli and contemporaneously in a number of basilicas on the Greek peninsula. It seems to be derived from the arrangement of pastophories flanking the forechoir, a pattern common throughout the Near East. Thus S. Pietro in Vincoli is among the earliest examples of a Near Eastern influence on Early Christian architecture in Rome.

S. PIETRO in Vincoli is known to every pilgrim to Rome if only because it houses Michelangelo's statue of Moses. The church lies at the S.W. corner of the Esquiline Hill, which slopes steeply down to the valley, at the bottom of which, in ancient

---

[1] The main results of this article formed part of a paper read at the Annual Meeting of the American Philosophical Society held in Philadelphia in April, 1939. The research on which it is based was done in connection with the preparation of the second volume of the *Corpus Basilicarum Christianarum Romae*, published by the Pontifical Institute for Christian Archæology in Rome and supported by a Grant from the Penrose Fund of the American Philosophical Society during the years 1938 and 1939.

The drawings used in this article were done by Mr. Wolfgang Frankl, architect, to whom I wish to express my sincere thanks.

times, ran the "Suburra", the famous street which connected the region of the Fora with the eastern height of the Esquiline. A steep staircase, protected by a XIII century tower, leads from the valley to a large plaza which extends in front of the church. From there four steps lead up to the porch of the basilica. This porch is a two-storied building with five arches on octagonal piers on the ground floor and with five rectangular windows on the second floor. In its interior the church presents itself as a basilica with axis running from west-southwest to east-northeast[2] (Fig. 1). The nave is separated from the aisles by two arcades of eleven semicircular arches each, which are carried by ten Doric columns. The aisles, without windows, are covered by groin vaults. The wooden barrel vault of the nave is segment shaped and is interpenetrated by three lunettes which contain three rectangular windows to the north and three sham windows to the south. The nave is followed by a transept covered with three groin vaults and this in turn by three semicircular apses. To the north the church is accompanied by the buildings of the present convent, a long narrow structure which reaches from the front line of the porch to the beginning of the transept; to the south it is enclosed by the buildings of the R. Scuola di Ingegneria which contain, among other remnants, the well-known cloister of S. Pietro in Vincoli.[3] The S. wing of the transept is skirted by a Medieval structure which contains the sacristy on the ground floor and some other rooms on the upper floors. Behind the church extends a small courtyard and behind this the large gardens which contain the vast remains of the Thermae of Trajan. Only a very few Roman remnants have been found in the neighborhood of the church:[4] remains of a street pavement, more than 80 m long, which skirts the N. side of the plaza in front of the church and fragments of another street pavement, east of the apse, in the narrow street, the Via delle Sette Sale, which skirts the convent building north of the church. One would assume that they belong together and that the direction of the Via delle Sette Sale, including the curve which it makes just at the beginning of the transept, goes back

---

[2] Throughout this paper I shall use, for convenience, the simplified directions W. for W.S.W., N. for N.N.W., E. for E.N.E., and S. for S.S.E.

[3] P. M. Létarouilly, *Edifices de Rome Moderne*, II, Paris, 1856, Pl. 140 ff.; see also Bufalini, map of Rome, 1551 (ed. Ehrle, 1911).

[4] R. Lanciani, *Forma Urbis Romae*, Milan, 1896, Pl. 22 and 23.

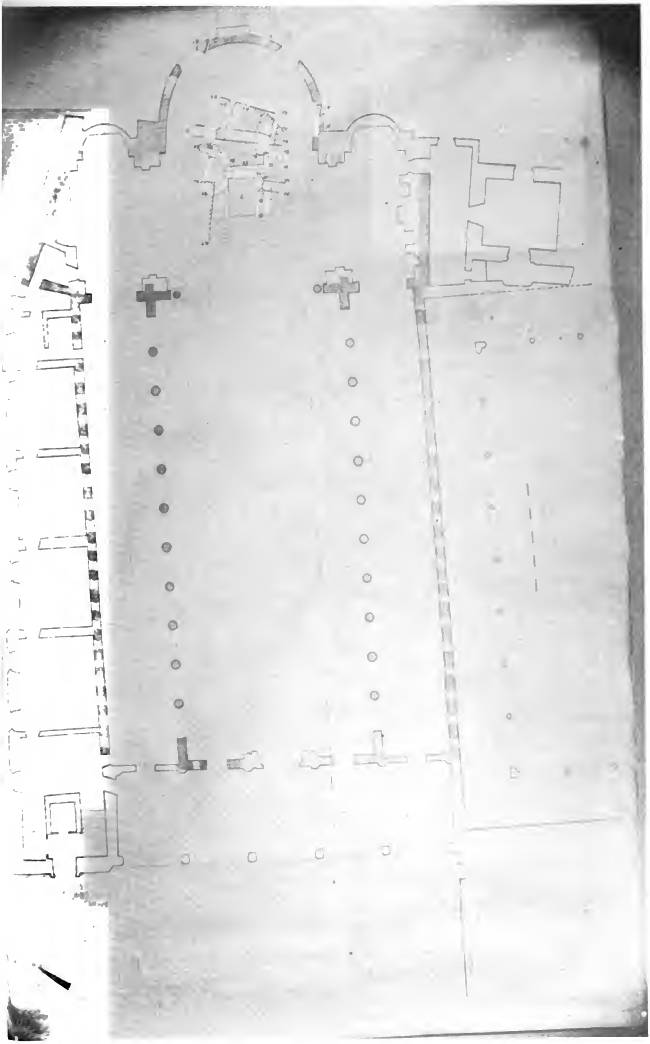


FIG. 1. S. Pietro in Vincoli, plan (drawing W. Frankl).

to Roman times. A few more traces of a pavement and two short pieces of walls have been found south-southeast of the church at a distance varying from 30 to 60 m.

It is obvious, of course, that the church in its present state contains a number of elements which point to later reconstructions (Fig. 2). Only a few years ago the triumphal arch was completely redecorated and given a classical moulding.[5] In 1876/77 the chancel was much restored: an open confessio was arranged in the middle of the transept, exactly in front of the High Altar, and the XV century shrine, containing the chains of Saint Peter, was placed in this confessio.[6] The High Altar itself, reconsecrated in 1877,[7] was rebuilt at the very borderline of apse and transept and placed under a highly decorative canopy. Below the High Altar a small crypt was constructed for the relics of the Maccabean Brothers which were rediscovered at that time. Simultaneously, a new pavement was laid in the

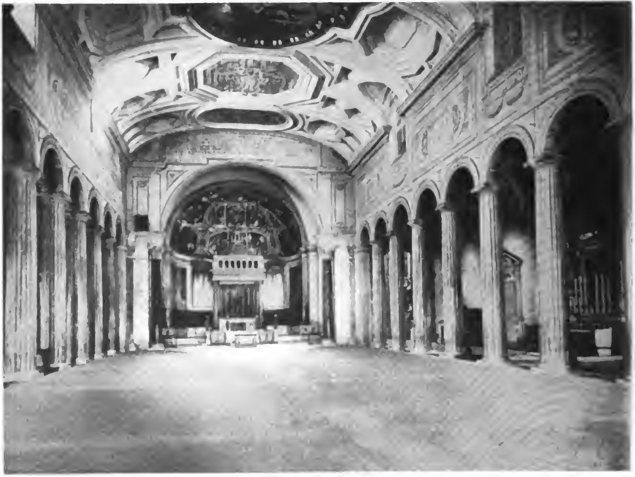


FIG. 2. S. Pietro in Vincoli, interior.

---

[5] Compare the old Alinari photo 6198 with the new photograph, Fig. 2.

[6] P. Mencacci, *Brevi notizie sulle Catene di San Pietro*, Rome, 1877, passim.

[7] Inscription in the confessio; see Mencacci, *op. cit.*, p. 67.

chancel[8] and a number of repairs were executed above the vaults of the transept. The upper part of the triumphal arch was restored, as appears from its XIX century masonry, and a door was arranged between the roof of the N. transept and the upper floor of the convent (see Fig. 12, right edge). Smaller repairs were executed over the whole exterior of the main apse (Fig. 3), including the brick archivolts of its windows. Before that time the church had a very different aspect which can be reconstructed from older plans and reproductions (Fig. 1, 4).[9] The altar stood further back in the apse and was raised on a rectangular platform. Two steps, instead of three, led from the

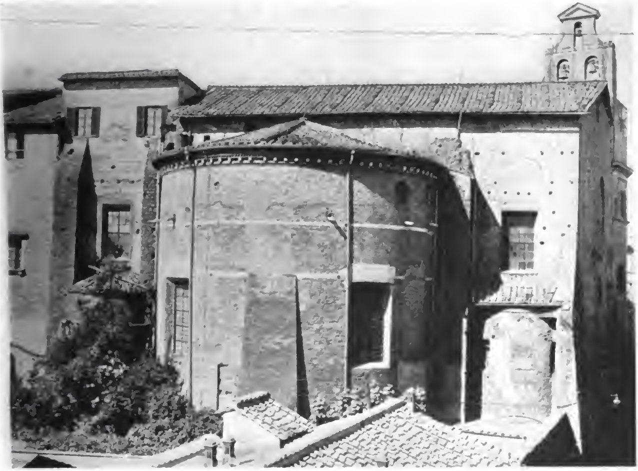


FIG. 3. S. Pietro in Vincoli, apse, exterior.

---

[8] Mencacci, *op. cit.*, p. 66 f.

[9] H. Huebsch, *Die altchristlichen Kirchen* . . . , Carlsruhe, 1862/63, Pl. IX, 7, 8; C. J. Bunsen, J. G. Gutensohn, J. M. Knapp, *Die Basiliken des christlichen Roms*, Munich, 1843, Pl. VIII, XI; L. Canina, *Ricerche sull' Architettura . . . dei Tempi Cristiana* . . . , Rome, 1846, Pl. LVI. Aside from the ground plans and sections contained in these works, an interior view is given in: L. Rossini, *I monumenti piu interessanti di Roma*, n. d. (but about 1818). An unpublished drawing, the groundplan of the old chancel and the remnants found below the apse in 1876, is preserved in the Lascito Lanciani XXXIX, Vol. 1, f. 49, no. 30930, at the R. Istituto di Storia dell' Arte e d' Archeologia, Rome. See below, n. 111.

transept to the apse. There existed, of course, no confessio, but, in front of the apse, a choir screen, with a wide opening in its center, projected to about the middle of the transept.

Earlier repairs had been made in the late XVIII and early XIX centuries. The bases of the Doric columns must have been added to the shafts about that time; they are certainly not original. They are made of Carrara marble, not Parian like the columns, and instead of standing under the shafts, they are added to them in two half rings, a procedure which was common in Rome during the XVIII century whenever bases were to be added to older columns.[10] In 1765 the pavement of the nave was laid,[11] and the portal was flanked by two columns and sur-

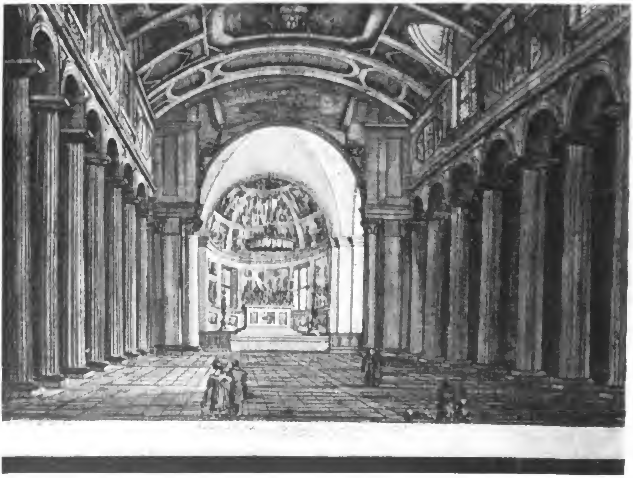


FIG. 4. S. Pietro in Vincoli, interior (engraving L. Rossini).

---

[10] The device was used, for example, at S. Maria Maggiore in 1743–50. I have not been able to establish the exact date at which the bases were added to the columns at S. Pietro in Vincoli. Seroux d'Agincourt, who wrote his book (J. B. L. G. Seroux d'Agincourt, *Histoire de l'Art . . .*, Paris, 1823) between 1781 and 1789 (*Ibidem*, I, p. 6 ff.), saw the columns still without bases (*Ibidem*, III, p. 16; IV, Pl. XXI, Figs. 2 and 6), while Rossini's engraving, which was done about 1818, already shows them.

[11] Inscription inside and over the main portal; V. Forcella, *Iscrizioni delle chiese . . . di Roma . . .*, IV, Rome, 1869–84, p. 91.

mounted by a tablet, with a related inscription, surrounded by a splendid aedicula. The segmental barrel vault was erected over the nave in 1705;[12] an inscription at the triumphal arch,[13] dated 1706, expressed the thanks of the abbot and his congregation to Prince G. B. Panfili for having the ceiling constructed and to Cardinal Marcello Durazzo, title cardinal of the church from 1686 to 1710,[14] for having it painted. The contemporary guide books elaborate on this statement. The ceiling was designed by Francesco Fontana,[15] while the ceiling painting was executed by G. B. Parodi from Genova.[16] At the same time the paintings in the apse were renovated by Giacomo Carboni,[17] the right lateral apse was restored,[18] the columns of the nave were done over,[19] and three new windows, those to the south being sham windows, were arranged on either side of the clerestorey.[20] Almost simultaneously, the ceilings in the portico along the facade were redone by Cardinal Giovanni Antonio Davia [21] and grills were made [22] bearing the coat of arms of Clement XI (1700–1721). For stylistic reasons the decoration of the triumphal arch, as it is still visible in older photographs, the decorations of the upper walls of the nave, the mouldings of the arches above the columns and, particularly, the Corinthian capitals of the columns at the triumphal arch and of the pilasters beside these, and the four windows right and left of the main door [23] must also belong to this same period.

---

[12] The date, 1750, which is given by A. Nibby, *Roma nel MDCCCXXXVIII*, I, pt. II, p. 669, is evidently a misprint.

[13] It was entirely changed some time ago, but the original text is preserved by Forcella, *op. cit.*, IV, p. 91 and in an old photograph, Alinari, no. 6198.

[14] Ch. Berton, ed. J. P. Migne, *Dictionnaire des Cardinaux*, Paris, 1857, col. 870 (= J. P. Migne, *Troisième et dernière encyclopédie théologique*, Vol. 31).

[15] Francesco Fontana, son of Carlo Fontana, 1668–1708; see Thieme-Becker, *Künstlerlexikon*, XII, Leipsic, 1916, p. 178, and L. Pascoli, *Vite de' pittori, scultori ed architetti*, II, Rome, 1730–36, p. 545 ff. N. Roisecco, *Roma antica e moderna*, II, Rome, 1765, p. 481, gives by mistake the name of the architect as Carlo instead of Francesco Fontana.

[16] O. Panciroli, F. Posterla, F. Cecconi, *Roma Sacra e Moderna*, Rome, 1725; p. 94; Roisecco, *op. cit.*, p. 481 f.; Pascoli, *op. cit.*, II, p. 547; F. Titi, *Descrizione delle Pitture . . .* , Rome, 1763, p. 478.

[17] Titi, *l. c.;* Nibby, *op. cit.*, p. 669.

[18] Titi, *op. cit.*, p. 240.

[19] Roisecco, *op. cit.*, p. 482; Nibby, *op. cit.*, p. 669.

[20] *Ibidem.*

[21] *Ibidem;* according to Berton, *op. cit.*, col. 787, he was Cardinal of S. Callisto from 1713 to 1737 and Cardinal of S. Lorenzo in Lucina from 1737 to 1740.

[22] Nibby, *op. cit.*, p. 669; Roisecco, *op. cit.*, p. 482.

[23] One sees clearly that they are older than the two columns of the main portal.

An earlier period of activity, though on a considerably smaller scale, had taken place late in the XVI century.[24] Cardinal Antonio Granvella, who held the title from 1570 to 78, had a second storey built over the portico, as it appears from the inscription, "ANT. CAR. GRANVELANUS", on the architrave of its windows.[25] On July 15, 1580 all the altars were reconsecrated by Gregory XIII.[26] The main restoration of that period, however, was the redecoration of the apse, in 1577, by Jacopo Coppi from Florence (Fig. 2).[27] Simultaneously, the two wide rectangular windows were arranged which light the apse at present. Two more rectangular windows, similar in shape but smaller, were disposed in the center of each of the lateral apses. A broad low buttress was erected against the center of the main apse. At the same time, two long rectangular windows of the type used in the main apse were pierced through the E. wall of the transept above the lateral chapels.[28] They are surmounted by wooden architraves with segment shaped relieving arches above them. One would assume that similar arches, instead of the XIX century brick architraves, existed above the rectangular windows of the main apse, and that also the three windows in the N. clerestorey of the nave, which are mentioned by Mellini as having existed in the middle of the XVII century,[29] had been arranged about 1575. Likewise, the small bell tower over the N.W. corner of the N. transept must have been executed during this rebuilding period, since its forms definitely point to the late XVI century (Fig. 3).

---

[24] The XVII century repairs were insignificant: C. B. Piazza, *La Gerarchia Cardinalizia*, Rome, 1723, p. 512, mentions repairs of the roof in 1676.

[25] Ugonio, Cod. Barb. lat. 2160 (Rome, Vatican Library), f. 172: "Ne il Cardl Presente Granvelano ha mancato di farvi fabriche et sue sono quelle sopra il portico della chiesa".

[26] Cod. Vat. lat. 11911 (Rome, Vatican Library); I am not quite certain whether this date is correct, and, unfortunately, I am unable to check it at present. Consecrations usually take place on Sundays and July 15, 1580 was not a Sunday; see H. Grotefend, *Zeitrechnung des deutschen Mittelalters und der Neuzeit*, I, Hannover, 1891–98, Pl. II and III.

[27] Roisecco, *op. cit.*, p. 480; Titi, *op. cit.*, p. 240; Ugonio, Cod. Barb. Lat. 2160 (Rome, Vatican Library), f. 172, was not quite sure of the date. "La tribuna e stata fatta dipingere l'anno 157 . dal R. P. . . . il 3º Anno del suo Generalato."

[28] Mellini, Cod. Vat. lat. 11905 (Rome, Vatican Library), f. 125 ff.: "La nave trasversa . . . piglia il lume a oriente da quattro finestre quadre . . . due sopra le cappelle e due nel semicircolo della Tribuna. . . ."

[29] Mellini, *l. c.*: ". . . questa nave piglia il lume ad occidente da una finestra tonda a settentrione da tre finestre quadre a mezzogiorno le finestre son finte. . . ."

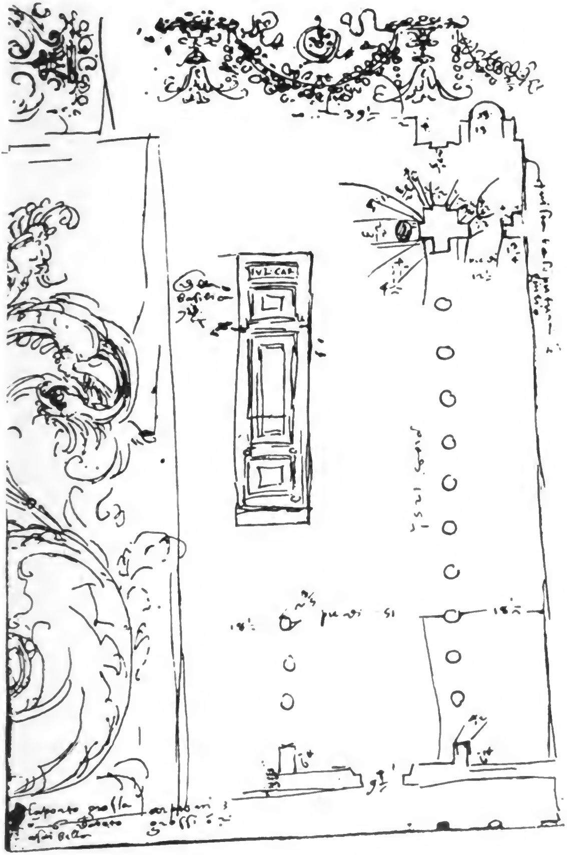


FIG. 5. S. Pietro in Vincoli, plan (drawing G. A. Colonna da Tivoli).

It is worthwhile, I think, to attempt a reconstruction of the aspect of the church before these repairs, that is, at about the time when, from 1542 to 1545, Michelangelo erected the tomb of Julius II against the S. wall of the S. transept. As today, the building consisted of a nave, accompanied by two aisles, of a transept and of three apses which opened onto the transept. The nave had an open timber roof with beams that rested on wooden leaf consoles; both beams and consoles are preserved above the XVIII century vault. The aisles were covered with the strongly domical groin vaults which still exist; they have no separating transverse bands and are strengthened by tie-rods. Three similar groin vaults cover the transept, a square one over the center bay and two oblong ones over the wings. The cross piers, of a somewhat irregular shape, which separate the nave from the transept, and likewise the corresponding piers between the main apse and the lateral apses existed at that time. This is proved by two groundplans of the XVI century, one by Antonio da Sangallo, the Younger (1486–1546)[30] and the other, with exact measurements, by Giovanni Antonio Colonna da Tivoli (1554)[31] (Fig. 5). The straight wall, which, in the interior, straightens the irregular shape of the N. transept, must also have existed at that time. It contains the niche in which the shrine for the chains of Saint Peter used to stand, and it must be contemporaneous with the transept vault that rests on it and above which it does not continue. The main apse, covered with a semidome, contained the old High Altar raised on a platform. The apse itself was raised two steps above the transept as were the lateral apses, where these steps are still preserved. At this period oblong windows with round arches were arranged in the main apse and in the center of the lateral apses; traces of two of the windows, filled with rubble work, are still visible next to the XVI century windows of the main apse and above the XVI century windows of the lateral apses (Fig. 3).

The N. transept wing sheltered the bronze shrine for the chains of Saint Peter,[32] the S. wing, the tomb of Julius II. As today, a portico extended along the facade. It opened in five semicircular arches, carried by four octagonal piers in the middle and by two half piers at the corners, all with palmette capitals. Six half octagonal piers, which lean against the wall of the facade proper, prove that, before the XVIII century, the portico was covered with groin vaults. A number of drawings show that this portico was one storied, and that it had a lean-to roof.[33] Other drawings help to further clarify the exterior aspect of the church. Above the roof of the portico, the facade seems to have contained a number of long arched windows, possibly three, and an oculus in the triangular gable.[34] Similar arched windows opened in the clerestory.[35] A drawing of about 1540 [36] shows that the long round arched windows of the apse were filled with a two-light tracery and that, aside from the two windows, traces of which are still visible next to the present rectangular XVI century windows, one more window existed exactly in the middle of the apse. Its site was evidently hidden when the thick low buttress was erected against the middle of the apse. When this drawing was made, a small bell tower rose from a tall building to the south of the transept; yet as late as the second quarter of the century,[37] a regular campanile seems to have risen from the S. transept of the church proper. Along the N. aisle there extended, according to this same drawing, the palace of S. Pietro in Vincoli, the remnants of which are preserved in what is the present convent. It had two stories, each with five rectangular windows with cross mullions, and a third story with a loggia; the whole facade was covered with a decoration of stucco or graffito work. On the opposite side of the church extended what then was the convent proper (and what at present is the Scuola di Ingegneria) and its cloister.

---

[30] Florence, Uffizi, Dis. arch. 1160 (A. Bartoli, *I Monumenti antichi nei disegni degli Uffizi . . .*, III, Rome, 1914–22, Pl. CCLXXXIV, Fig. 470).

[31] Cod. Vat. lat. 7721, formerly sp. Colonna 7721 (Rome, Vatican Library), f. 20′; unpublished. The measurements are given in Roman feet.

[32] It is mentioned as in this place from the late XVI century on: Ugonio, Cod. Barb. lat. 2160 (Rome, Vatican Library), f. 172; Mellini, Cod. Vat. lat. 11905 (Rome, Vatican Library). It was still there in the early XVIII century (Titi, *op. cit.*, 1763, but perhaps just taken over from the edition of 1736) but was certainly removed to the sacristy before 1765 (Roisecco, *op. cit.;* Nibby, *op. cit.*) where it remained until 1876.

[33] Map of Rome, 1474, Laur. Red. 77 (Rome, Vatican Library), (G. B. De Rossi, *Piante iconografiche e prospettiche di Roma*, Rome, 1879, Pl. IV); Map of Mantova, 1534, probably from a prototype of before 1483 (De Rossi, *op. cit.*, Pl. VII); Anonymous, View of Rome, about 1490, Cod. Escurial., f. 40 (H. Egger, *Römische Veduten*, II, Vienna, 1931, Pl. 104, and idem, *Codex Escurialensis, Sonderdruck des österreichischen archäologischen Instituts*, Band IV, Text, Vienna, 1906, Pl. IV); Dosio, panorama of Rome, about 1560, Florence, Uffizi, Dis. arch. 2523 (Bartoli, *op. cit.*, V, Pl. CDXII, Fig. 748). The drawings of 1532–36 in Heemskerk's sketchbook at Berlin, Kupferstichkabinett, f. 55 and f. 92 (H. Egger and Ch. Hülsen, *Die römischen Skizzenbücher des Marten van Heemskerk*, II, Text, Berlin, 1913 ff., Pl. 74, 121, pp. 34, 51), are small sketches and not very reliable.

[34] Map of Mantova; Heemskerk; Dosio; see preceding note.

[35] Map of Mantova; see above, n. 33.

[36] French or Piedimontese Anonymous, before 1551; formerly Rome, Coll. L. Pollak (Egger, *Römische Veduten*, II, Pl. 148).

[37] Heemskerk, see above, n. 33; it appears also in Wyngaerde's panorama, which is usually dated about 1550 (Egger, *Römische Veduten*, II, Pl. 108).

The constructions at the church of S. Pietro in Vincoli proper, such as the top parts of both nave and transept walls which carry the beams of the XV century roof and which, therefore, must be contemporaneous with it, are done in a very characteristic masonry of roughly hewn tufa stones (Fig. 12). It is a type of construction which frequently appears in Roman buildings of the late XV century.[38] This same masonry was used for constructing the E. wall of the S. transept, the two lateral apses, and the buttresses which bond with these apses, one on either side of the S. apse and one more north of the N. apse. This last buttress interpenetrates an inclined wall which abuts the N. wall of the N. transept. All these parts must have been added to the building in the late XV century (Fig. 3).

This observation is particularly important for the two lateral apses. Including the foundation walls of the S. apse, they are done in the same masonry and have no structural connection whatsoever with either the main apse or the walls of the transept. This is confirmed by the fact that the foundation walls of the S. apse are at a level about 50 cm above that of the main apse.[39] Thus the lateral apses should be considerably later than the main apse (Fig. 6).

This whole lay-out, particularly the arrangement of the church, follows a pattern which was common in Rome during the last third of the XV century. It has not, however, received sufficient consideration. It is the pattern used under Sixtus IV, Innocent VIII, and Alexander VI for the restoration of Early Christian and Medieval basilicas. Whether one studies S. Cecilia or S. Agnese f. l. m., S. Eusebio or S. Maria Maggiore, S. Marco or SS. Apostoli, S. Agata dei Goti or S. Croce in Gerusalemme,[40] one finds the same features: the open timber roof or the flat ceiling in the nave, groin vaults in the aisles and in the transept, if the church has a transept, a groin vaulted portico in front of the building, the cloister to one side. Occasionally even new churches were constructed on this pattern, which was, evidently, the XV century conception of the appearance of an Early Christian church. Such a "neo-Early Christian" basilica of the late XV century is S. Lorenzo in Damaso in the Cancelleria. The prototype of the whole group seems to have been the plan which was followed when Nicolaus V attempted to rebuild St. Peter's from 1451 to 1455. Whether Bernardo Rossellino or Leone Battista Alberti was responsible for the plan, or whether it had been started before they came to Rome in December, 1451 and in 1452, respectively, is of little importance here.[41] What is significant is that Manetti, in his life of Nicolaus V,[42] clearly indicates that the intention was to construct groin vaults in the aisles and in the transept of St. Peter's, and that, likewise, the narthex in front of the church was to be vaulted. As in all the churches mentioned, the nave was to have an open timber roof or a flat ceiling. The aspect of S. Pietro in Vincoli in the XV century fits perfectly into the pattern followed in rebuilding Early Christian basilicas during this period.

---

[38] For example, the campanile of S. Agnese f. l. m., dated 1479.

[39] The foundations of the N. apse are covered by the pavement of the courtyard.

[40] S. Agata dei Goti, restoration of 1461–69: vaults of aisles, portico existed already; S. Agnese f. l. m., restoration of probably about 1479: vaults of aisles; SS. Apostoli, restoration of 1453–77: repairs, flat ceiling of nave, portico, cloister; S. Cecilia, restoration of 1484 (?): vaults of aisles, portico existed already; S. Croce, restoration of 1460–93: flat ceiling of nave, vaults of aisles and of transept; S. Eusebio, restoration of 1447–55: flat ceiling of nave, vaults of aisles; S. Marco, restoration of 1455–71 (?): flat ceiling of nave, vaults of aisles, portico; S. Maria Maggiore, restoration of 1450 ff.; flat ceiling of nave, vaults of aisles and transept, portico existed already; etc.

There are many documents to prove that this transformation of S. Pietro in Vincoli took place during the last third of the XV century. Contrary to the general opinion, however, this rebuilding consisted not of one, but of a series of restorations. Nicolaus of Cues, who held the title of cardinal from 1448 to his death in 1464, had started the reconstruction, although on a small scale. He began repairs on the roof,[43] erected an altar, with two porphyry columns, in the N. transept,[44] near which his tomb was erected in 1465, and bequeathed a sum of money for the repair of the church.[45] The new title cardinal, created in 1467, was Francesco delle Rovere, who, in 1471, became Sixtus IV. While cardinal, he first repaired the palace of S. Pietro in Vincoli, and these repairs were still going on in the fall of 1471, after he had become pope.[46] Also, the roof of the church was restored in, or shortly before, 1471.[47] Yet the transformation of the church proper was due to the pope's nephew, the new title cardinal Giuliano delle Rovere (1471–1503). The vaults of both the transept and the main portal bear the papal coat of arms of the delle Rovere, and, since the architrave of the door used to bear an inscription which can have referred only to Giuliano delle Rovere, they must thus be dated after 1471.[48] The wooden door of the nave also carries an inscription referring to Giuliano (Fig. 5). The bronze shrine for the chains of St. Peter, which, before the XVIII century, stood in the left transept, is dated 1477 and bears, as well as their coat of arms, an inscription referring to Sixtus IV and Giuliano delle Rovere.[49] The portico also must have been erected before 1475 or at least before 1483 since, from this time on, it appears on the maps of Rome.[50] Vasari had attributed the remodelling of the church as well as the erection of the portico to Baccio Pintelli.[51] The date of the cardinal's palace, the building of which Vasari attributed to Giuliano da Sangallo,[52] cannot be definitely ascertained.[53] Since, however, the portico leans against it, the palace must have been started contemporaneously with, or even earlier than, the portico.[54] Likewise the convent to the south, including the cloister, must have been started about this period, as is proved by the inscription "JUL. CAR. S. PET. AD VIN" on the architrave of one of its doors.[55] The name of the architect is unknown.[56] The inscription would point to a date before

---

[41] G. Dehio, "Die Bauprojekte Nikolaus V. und L. B. Alberti," in: *Kunsthistorische Aufsätze*, Munich and Berlin, 1914, p. 163 ff.

[42] G. Manetti, "*Vita Nicolai Quinti*," in: L. A. Muratori, *Rer. Ital. Script.*, Milan, 1723–51, III, 2, col. 934 ff.

[43] Mellini, Cod. Vat. lat. 11905 (Rome, Vatican Library), f. 137: Ugonio, Cod. Barb. lat. 2160 (Rome, Vatican Library), speaks more generally of repairs of the church.

[44] F. Martinelli, *Roma ex ethnica sacra*, Rome, 1653, p. 284 f.

[45] E. Müntz, "Les arts à la cour des papes," III; in: *Bibliothèques des Ecoles françaises d'Athènes et de Rome*, fasc. 28, Paris, 1882, p. 165: On Jan. 15, 1472, an emerald, worth 70 ducats, is sent by Sixtus IV to the new title cardinal, Giuliano delle Rovere, "pro residuo satisfactionis certarum pecuniarum per bo. me. dominum Nicolaum de Cusza . . . legatum pro reparatione ecclesiae. . . ."

---

[46] Müntz, *op. cit.*, p. 164.

[47] Müntz, *l. c.*

[48] Nibby, *op. cit.*, p. 665: "Astra Palatinis quae tangit ab aedibus hospes. Hac primum nata est Julia quercus humo".

[49] A. Venturi, *Storia dell' arte italiana*, VI, 1908, p. 928; it is irrevelant for our problem whether the shrine was executed by the Pollaiuoli or by Caradosso.

[50] See above, n. 33.

[51] See also E. Lavagnino, "L'architetto di Sisto IV", in: *L'Arte*, XXVII (1924), p. 4 ff. It is, however, doubtful whether Pintelli came to Rome before 1482.

[52] Vasari, *op. cit.*, IV, p. 278 f. dated it after 1492; this is quite impossible since Giuliano delle Rovere was absent from Rome from 1492 to 1503.

[53] H. L. Heydenreich in Thieme-Becker, *op. cit.*, XXIX, p. 406, dates it "1477(?)" without indicating his reasons.

[54] It seems that Giuliano delle Rovere did not complete the palace; see A. Donatus, *Roma vetus ac recens*, Amsterdam, 1695, p. 325.

[55] Letarouilly, *op. cit.*, II, Pl. 141, 142; text p. 320 ff.

[56] While A. Condivi, *Vita di Michelangelo Buonarotti*, Rome, 1551, ch. XXV (ed. Pisa, 1823, p. 25), had attributed the convent to Bramante, the general attribution is

1492, since Giuliano delle Rovere had to leave Rome in this year and could not return until after Alexander VI's death in 1503, when he himself was elected Pope Julius II. The cloister, however, was completed only later as other inscriptions in it refer to cardinal Sixtus delle Rovere, who held the title from 1507 to 1517.[57] It becomes evident that the building activity was long, extending over almost half a century, and was undertaken by the whole house of the delle Rovere, not by any single one of its members.[58]

II

If one deducts all the parts which were added to the church of S. Pietro in Vincoli from the XV to the XIX century, the edifice presents itself in its original aspect. With the exception of one minor change,[59] the building had preserved this original state up to the late XV century.

In the original church, as today, nave and aisles were separated by arcades, each consisting of eleven arches (Fig. 6).[60]

---

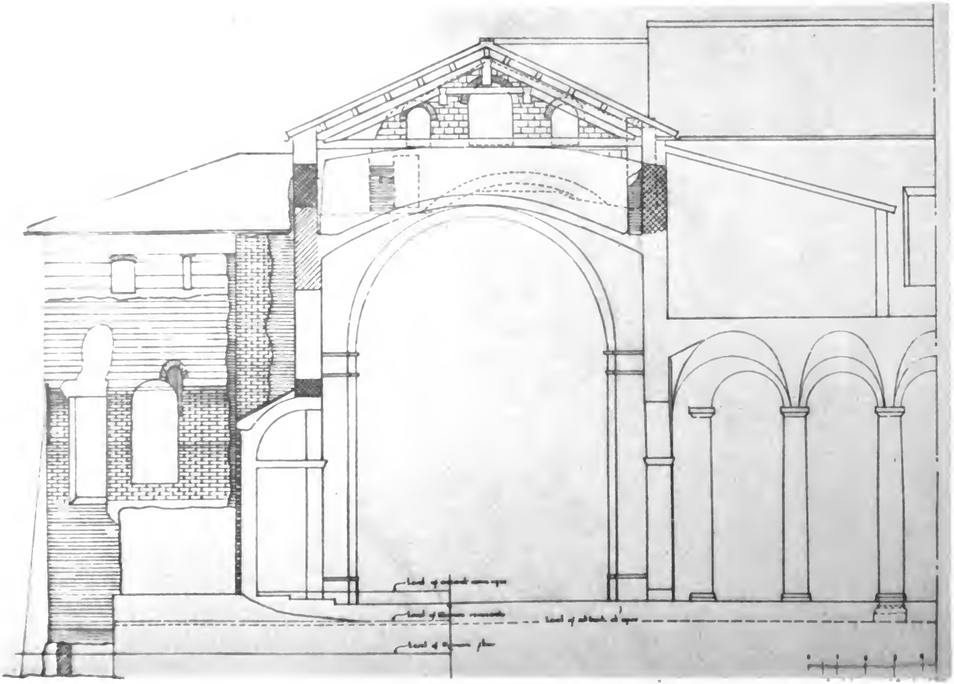
to Giuliano Sangallo, the Elder. This is also the opinion of G. Clausse, *Les Sangallo*, I, Paris, 1900, p. 148 ff. Clausse adds that the cloister was erected "dans les premiers jours de . . . 1490", but he does not give any reasons for this precise statement. The capitals of the columns in the cloister, according to Clausse, *l. c.*, would have been executed about 1503, partly before and partly after Julius' election as pope, because they bear the coat of arms of the delle Rovere alternatingly surmounted by tiaras and by cardinals' hats. This conclusion is not quite convincing since the coat of arms in this combination might as well allude to a period in which one delle Rovere was pope and another, cardinal. Such a constellation occurred between 1471 and 1484, under Sixtus IV and Giuliano delle Rovere, and between 1507 and 1517, under Julius II and Sixtus delle Rovere.

[57] Létarouilly, *op. cit.*, p. 320 ff., while the inscription on the well head in the cloister shows that this well was begun under Julius II (1503–13), the inscription on the colonnaded structure over the well refers to the cardinalate of Leonardo delle Rovere (1517–20).

[58] The erection of the organ loft in the N. transept of the church by Cardinal Alessandro Cesarini (Ugonio, Cod. Barb. lat. 2160 (Rome, Vatican Library), f. 172) before 1542 may be considered as having terminated this building activity. He also donated shortly after 1530 the wooden door of the sacristy; see A. Paolucci, "Di una porta lignea scolpita nella sagrestia di S. Pietro in Vincoli," in: *Roma*, X (1932), p. 23 ff.

[59] See below, p. 377 f.

[60] The measurements of the original building are:
Length: from facade to apse: clear, 59.75 m = 202 R. ft.; including walls, 61.35 m = 207½ R. ft.; center of wall to center of wall, 60.55 m = 205 R. ft.
: facade to triumphal arch: clear, 38.50 m = 130 R. ft.; including walls, 40 m = 135⅓ R. ft.; center of wall to center of wall, 39.40 m = 133⅓ R. ft.
Width: total: clear, 28.61 m = 97 R. ft.; including lateral walls, 30.25 m = 102½ R. ft.; center of wall to center of wall, 29.45 m = almost 100 R. ft.
: nave: clear, 15.80 m = 53½ R. ft.; including colonnades, 17.35 m = 58½



The arches were carried by ten Doric columns which, including their capitals, were 6.80 m high.[61] They are certainly pilfered from some ancient building,[62] yet, in spite of what has been said to the contrary, they are not absolutely homogeneous.[63] They differ in diameter; those to the south vary from 75 to 84 cm, those to the north from 80 to 86 cm, and they differ in the upper parts of their shafts where the flutings are sometimes more, sometimes less curved. It seems that the columns originally had bases. Of course, they did not have the present ones which are definitely of the late XVIII or early XIX century. Early in the XVI century Fra Giocondo made an accurate drawing of two bases from the church "a santo pietro in vinchullo Basamenti Doricij."[64] They are beautiful bases, each consisting of plinth and torus (Fig. 7); above the torus, at the foot of the shaft, is a fillet and then the curved beginning of the shaft proper. Their heights were different; one base was considerably lower and seems to have risen from a high socle which is clearly indicated in the Frate's drawing. The bases must have belonged to the columns of the arcades. Fra Giocondo says clearly that they are "Basamenti Doricij" and there are no other Doric columns in the edifice but those of the arcades. On the other hand, the two non-Doric columns which are in the church, the columns of the triumphal arch, still rise from their original bases which are attic bases with double tori. The measurement which Fra Giocondo indicated on the shaft of the column with the higher base, 1.195 m, and which must be half the circumference of the shaft, corresponds to a diameter of 0.76

R. ft; center of colonnade to center of colonnade, 16.60 m = 56 R. ft.
: aisles: clear, 5.70 m = 19½ R. ft.; including lateral wall and colonnade, 7.10 m = 23½ R. ft.; center of wall to center of colonnade, 6.50 m = 22 R. ft.

Distance of columns from axis to axis: 3.28 m = 11 R. ft., average. There are slight variations.

It is evident from these measurements that the measurements for the original building were taken, as a rule, from center of wall to center of wall since these respective measurements are the only ones which result in round figures.

[61] Seroux d'Agincourt, *op. cit.*, IV, Pl. XXI, Figs. 2 and 6, gives their height as 19 Parisian feet, 10 pouces = 6.44 m; Hübsch, *op. cit.*, gives it as 7.65 m!

[62] Their provenance cannot be ascertained.

[63] F. de' Ficoroni, *Le vestigia . . . di Roma antica*, Rome, 1744, p. 101, mentions that only eight of the columns are of Parian marble.

[64] Florence, Uffici, Dis. arch. 1542 verso (Bartoli, *op. cit.*, I, Pl. XXXIV, Fig. 61). The plinth of one was 8 oncie = 19.8 cm, of the other 6 oncie, 1 minuta = 15.1 cm high; the torus 8 oncie, 4 minute = 20.6 cm and 4 oncie = 9.9 cm high; the fillet 1 oncia, 1 minuta = 2.7 cm and 1 oncia = 2.5 cm respectively.

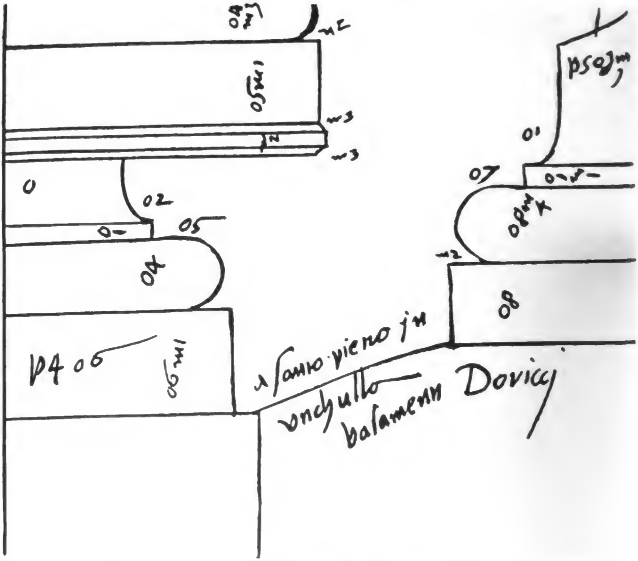


FIG. 7. S. Pietro in Vincoli, bases (drawing Fra Giocondo).

m;[65] the diameter of the other column was 0.866 m.[66] While this corresponds exactly to the diameters of the Doric columns (0.75–0.86 m), it far exceeds the diameter of the two columns at the triumphal arch which are only 0.68–0.70 m in diameter. Thus the Doric columns, or at least part of them, in all likelihood, were from the very outset placed on bases.[67]

The Fra Giocondo drawing, however, leads to another point; one base was 40.4 cm high, the other 25 cm high, and the lower one rose from a socle of unknown height. Including the fillet and the curve, which only about 5 cm higher merged into the straight part of the shaft, the shafts started 47 cm and 32 cm, plus the unknown height of the socle, respectively, from the original level. At present the columns show no trace of any curve or fillet in spite of the fact that the present bases cover only the lowest 27 cm of that shaft. This means that the columns must have been somewhat longer than they are at present[68] and that the original level of the nave must have been considerably lower than the present pavement. Where this original level was situated may be concluded from a number of other observations. First, in the wall of the main apse, there appears a setback, as frequently occurred in Early Christian basilicas, at the level of the pavement. This setback lies 70 cm below the present level of the nave. Second, when in 1876 excavations were undertaken in the apse, the walls of the buildings preceding the church were found at about 1 m below the level of

---

[65] 5 palmi, 3 oncie, 3 minute = 119.5 times 2 = 339 cm divided by π (3.14159) = 76 cm.

[66] Length of plinth, 4 palmi, 6 oncie = 1.04 m minus (projection of torus, 5 oncie =) 12.4 cm minus (curving of shaft, 2 oncie =) 5 cm = 0.866 m.

[67] This feature is unusual in classical Antiquity but it appears not infrequently in the Hellenistic architecture of Rome during the last centuries B.C. One may compare e.g. the columns of the temple at Cori with Doric capitals, slender shafts and bases with single torus. The lower thirds of their shafts are not fluted, and it is worth noticing that also Fra Giocondo does not indicate any fluting for the columns of S. Pietro in Vincoli.

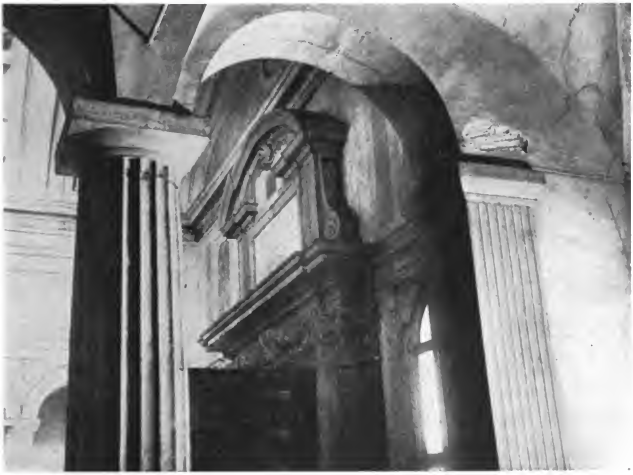


Fig. 8. S. Pietro in Vincoli, North arcade, detail of first arch.

---

[68] Ficoroni, *op. cit.*, p. 101.

the ancient chancel.[69] Since at that time the chancel was still raised about 34 cm above the nave, the level of these ancient walls would correspond to a level of 65 to 70 cm below the nave and thus approximately to the level of the setback mentioned. One has to imagine that the walls were razed just below the original level of the chancel when its pavement was laid (Fig. 6).

The arches of the arcade are old. An uncovering of their original structure has shown that they consist of long "roman" bricks set into broad mortarbeds, the technique used in the IV and V centuries. Their apexes are situated about 8.02 m above the present pavement (Fig. 8).

The clerestory above the arcade was originally 7.05 m high, before it was heightened in the XV century (Fig. 9). In the N. clerestory,[70] above each of the arches of the arcade, are traces of a window, terminated by a semicircular arch.[71] Originally there must have been eleven windows in each clerestory.[72] Both the width of the windows and the distances between them increase from the east to the west, from a cumulative distance of 3.15 m (10⅔ R. ft.) to one of 3.39 m (11½ R. ft.) from axis to axis of the windows. In height all the windows seem to be equal, 1.70 to 1.72 m from the sill to the springing of the arch and 2.62 m (9 R. ft.) up to the apex.[73] The arches of the windows spring from a small setback which is about 8 cm deep. The bricks of the voussoirs are 0.45 m (1½ R. ft.) long and are irregularly arranged, rather out of radius. At some later period some of the windows[74] were walled up with a masonry of opus mixtum. Otherwise the clerestory has preserved its original aspect and its original brick masonry. Remnants of the same masonry seem to be preserved also in the top part of the façade above the XVIII century ceiling.

---

[69] See Tomasetti's report in: G. B. De Rossi, *Bulletino di Archeologia Cristiana*, Ser. III, anno I, 1876, p. 73 ff.; see below, n. 103.

[70] Except for the spaces which are taken up by the rectangular XVI century windows and for the space above the easternmost bay which is hidden by a chapel on the top floor of the convent.

[71] The S. clerestory is hidden by the buildings of the R. Scuola di Ingegneria.

[72] They are from 1.77 m to 1.83 m wide, except for the westernmost window which is 2 m wide. The distance from one to the other window is: 1.38 m between the seventh and eighth, 1.43 m between the fourth and fifth and 1.48 m between the first and second.

[73] Hübsch, *op. cit.*, erroneously indicated the width of the windows as 1.95 m, their distances apart as 1.15 m. Rohault de Fleury gave them as 1.85 m and 1.83 m, respectively.

[74] Unfortunately we were not allowed to uncover systematically more than the westernmost window.

No trace is preserved of the original entrances to the basilica. At present there exists only the one doorway which leads from the portico to the nave and which, in its present shape, is of the XV century.[75] From the two XVI century plans by Antonio da Sangallo, the Younger, and Giovanni Antonio Colonna,[76] we know that this arrangement of only one door goes back at least to that period. Whether, before the XV century, three doors led

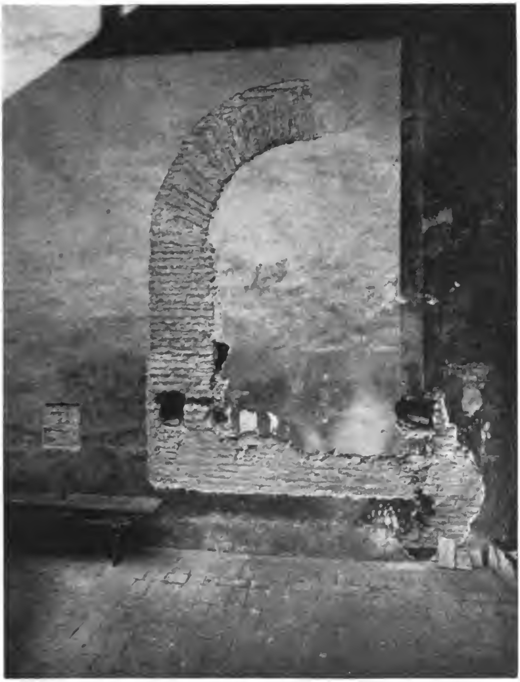


FIG. 9. S. Pietro in Vincoli, North clerestory, first window.

---

[75] See above, p. 366.

[76] See above, p. 362.

into the nave, analogous to the contemporaneous arrangement which existed at S. Lorenzo in Lucina and still exists at S. Maria Maggiore, cannot be ascertained at present. The aisles, at any rate, have had no doors, at least since the VII century as from the late VII century until 1683, when it was removed to its present place in the middle of the left aisle,[77] the altar of St. Sebastian, with its mosaic, was leaning against the W. wall of the N. aisle.[78] Consequently there could not be any entrance there. Nothing can be determined about the existence of an atrium in front of the church. The wide plaza, which still extends there, may very likely have replaced such an atrium; yet one will have to await the results of future excavations to ascertain this point.

The original aspect of the nave can thus quite definitely be determined. As a whole it does not depart in any way from the normal pattern of an Early Christian basilica in Rome. Much more interesting are the eastern parts of the edifice, the transept and the chancel.

The triumphal arch which leads from the nave into the transept rests on two beautiful porphyry columns with old attic bases.[79] Their Corinthian capitals seem to be made of stucco and belong probably to the XVIII century.[80] The arch itself evidently consists of two voussoirs; a break between them is clearly visible in older photographs.[81] However, it is unlikely that the lower arch was added later and that originally the triumphal arch was higher and wider. The few bricks of the lower arch, which are visible in the above mentioned photograph, look very much as though they were Early Christian and the break may easily have occurred between the two voussoirs of the *original* triumphal arch. The upper parts of the upper voussoirs, which are visible above the vaults of transept and nave, were repaired in 1876 as shown by their masonry.[82]

---

[77] The removal did not take place in 1577 as M. Armellini, *Le Chiese di Roma*, Rome, 1891, p. 208, assumed. At that date the altar was merely given a breve by Gregory XIII; see Forcella, *op. cit.*, IV, p. 83.

[78] Inscription of 1683 formerly at the altar, see Forcella, *op. cit.*, XIII, p. 425: "... REGULARES DTI MONASTERII .. NUPER EXPONI FECERUNT// IPSI ALTARE PRIMO DICTUM QUOD MURO PRAEFATAE EORUM ECCLESIAE PROPE ILLIUS JANUAM// ADHEREBAT ITA UT SACERDOTES AD ILLUD CELEBRANTES HUMEROS ALTARI MAJORI IPSIUS// ECCLESIAE OBVERSOS HABERENT INDE IN EIUSDEM ECCLESIAE NAVEM SEU ALAM LATERALEM A// SINISTRO INGREDIENTIUM TRANSTULERUNT IBIQUE DECENTIUS COLLOCAVERUNT. ..." For the date of the mosaic, which was possibly donated in connection with a pestilence that occurred in 672 (not in 680), see G. B. De Rossi, *Musaici Cristiani*, Rome, 1899, Pl. XX, and J. Wilpert, *Die Römischen Mosaiken . . .*, II, Freiburg, 1915, p. 1001 f. J. H. Parker, *Archeology of Rome, vol. XI, Church and altar decoration*, Oxford, 1876, p. 41, dated the mosaic erroneously in the XII century.

[79] See above, p. 369.

[80] See above, p. 359.

[81] Alinari, no. 6198.

The masonry of the transept, in those of its walls which are original, that is, in the whole of the N. wing, in the center part above the main apse and above the triumphal arch, and in the S. wall of the S. transept (Fig. 1), corresponds almost exactly to that of the clerestory of the nave. The E. wall of the S. transept has been completely restored with the tufa masonry which is characteristic of all the XV century repairs of the structure. In the original parts, the ratio of the bricks and mortarbeds is 8 to 8 or 8 to 7 every 50 cm; the height of the bricks varies from 3.2 to 3.9 cm and the mortarbeds are from 2.4 to 3.4 cm thick. There is not the slightest doubt that the upper walls of the N. transept and the wall of the N. clerestory are bonded to each other and are of one and the same building period (Fig. 6).[83]

The transept is completely irregular in plan (Fig. 1). Only its W. wall forms a right angle with the axes of the nave and of the aisles. The S. wall diverges slightly towards the south, and the N. wall converges markedly in the same direction. The N. wall is almost 2 m longer than the S. wall. Thus the E. wall, which in the N. wing runs parallel to the W. wall, has to slope down in order to meet the E. end of the S. wall. Although the E. wall of the transept has been considerably restored, its main direction, and thus the shape of the transept, are evident. It was rather narrow at the S. end and wide at the N. end, and while its S. wall projects 20 cm beyond the lateral wall of the S. aisle, the N. wall projects 43 cm beyond the N. aisle. While the W. wall and the N. end of the E. wall of the transept conform to the axial system of the nave, both the S. wall and the N. wall of the transept follow two entirely different axial networks. Also the thickness of the walls varies; the N. wall of the transept is only 0.58 m thick, while all the other walls are 0.78 m thick.

---

[82] See above, p. 357.

[83] I am sorry to have to disagree on this point with J. P. Kirsch, who, in several articles, "Die Entwicklung des Bautypus der altchristlichen römischen Basilika", in: *Römische Quartalschrift*, 43 (1935), p. 1 ff., and "Das Querschiff in den . . . Basiliken . . ." in: *Pisciculi, Studien . . . F. J. Dölger dargebracht,* Münster, 1939, p. 148 ff., has maintained that the transept is a later addition. E. Weigand in: *Byzantinische Zeitschrift*, 39 (1939), p. 569, follows Kirsch in this matter.

One window opened in the upper part of each of the W. and of the E. walls of the N. transept and two windows in its N. wall (Fig. 10). All are larger than the clerestory windows.[84] No traces of any windows are visible in the walls of the S. transept which are all heavily plastered. In the lower parts of the transept only one trace of an opening is visible, an arch, the top of which emerges at the exterior just above the lean-to roof over the N. lateral apse. The apex of its extrados (the intrados is not visible) is situated precisely 7.75 m above the present level, and consequently about 8.45 m above the original level of the church (Fig. 3). The arch is thus far too high for a doorway; it can have covered only a window. Its radius would have

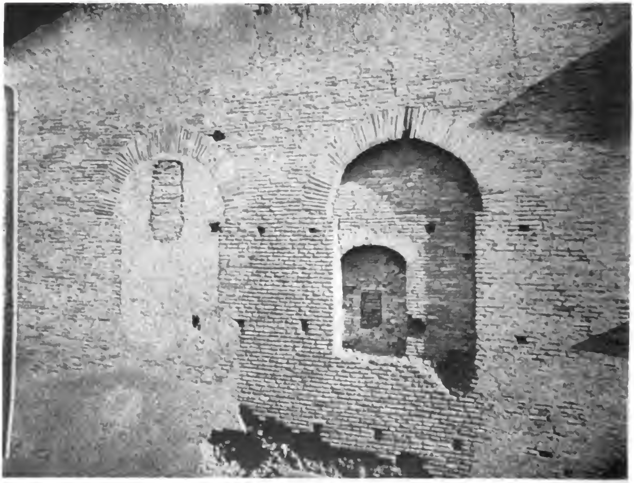


FIG. 10. S. Pietro in Vincoli, North wall of transept.

---

[84] They start 0.55 m below their level and end 0.10 m above it. Those in the E. and W. walls are 2.35 m wide (8 R. ft.), 2.20 m high (7½ R. ft.) up to the springing of the arch, and 3.35 m high (almost 11½ R. ft.) up to its apex. The other ones in the N. wall are equally high, but their width is only 2.20 m (7½ R. ft.) and 2.30 m (almost 8 R. ft.) respectively, the eastern one being somewhat larger.

corresponded to that of the window in the upper section of the wall. Nothing can be said about the existence of a similar arrangement in the S. wing, since the wall there is of the XV century.

Far more interesting than these windows are some other remnants which are preserved in the interior of the transept above the XV century vaults. There is a wide arch which spans the width of the transept between its center bay and its S. wing (Figs. 6, 11). Three window-like openings are arranged in a wall which is supported by this arch.[85] The arch, as well as the wall above, is painted in a XIV or early XV century pattern. The voussoirs of the arch consist of well-hewn travertine blocks; the masonry of the wall is brickwork, with 10 bricks and 9 to 10 mortarbeds every 50 cm.[86] It was heightened and its openings were closed with rubble masonry and with late brickwork only when, simultaneously with the erection of the vaults, the

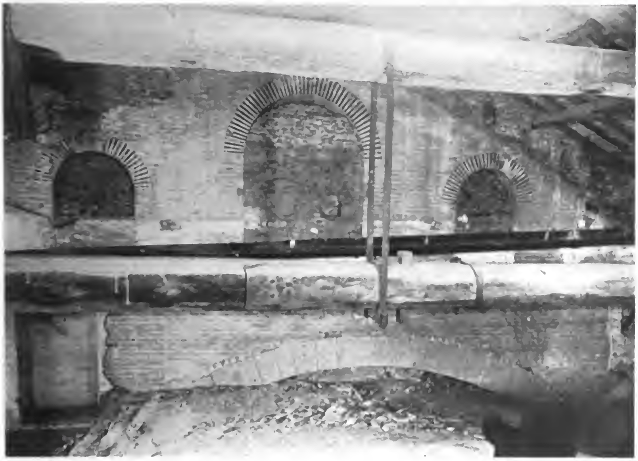


FIG. 11. S. Pietro in Vincoli, transept above vaults, view from North.

---

[85] The small rectangular door at the left was evidently cut into this wall when the XV century vaults were erected over the transept.

[86] The brick masonry shown on the photograph is painted. The genuine brickwork of the wall shows at the left, at the right jamb of the rectangular door.

present roof was constructed. The arch proper is certainly a late Medieval construction. It is an addition to the original edifice, evidently erected to support the contemporaneous Campanile which rose from this transept wing.[87] But the arch is not entirely Medieval; at the right and left ends it rests on two piers which protrude from the W. and E. walls, respectively, of the transept. Two corresponding piers protrude from the walls of the transept between the N. wing and its center bay. They are well preserved above the XV century vaults (Fig. 12) and, except for the topmost bricks, they are still as high as the top of the original transept walls. Their lower parts must be contained within the present crossing piers and within the piers between the main apse and lateral apses. Above the XV century vaults these pilasters are each 0.80 m wide, but they differ in depth. The pier to the northwest protrudes 62 cm from the W. wall of the transept, the one to the southwest, 45 cm; the pier to the southeast protrudes 88 cm from the E. wall. No exact measurements can be given for the pier to the northeast,

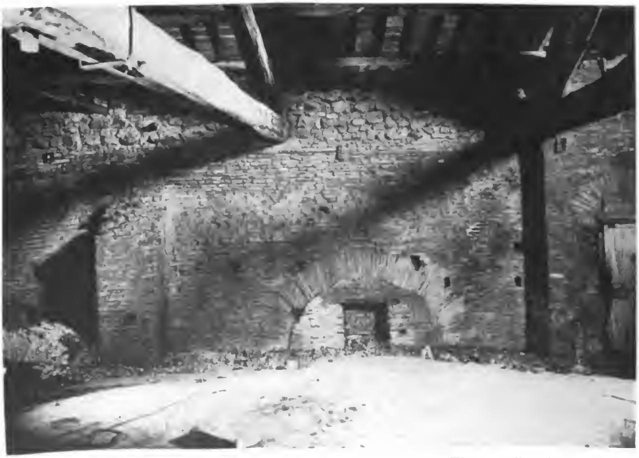


FIG. 12. S. Pietro in Vincoli, transept above vaults, West wall, pier between North wing and center bay.

---

[87] See above, n. 37.

since all but the very bricks next to the wall have been demolished. Yet the pier could not have protruded more than 1.00 m since it could not have been larger than the pier in the transept below, between the N apse and the main apse, within which its lower part must still be contained. All these piers have exactly the same masonry as the brick walls of the transept, and all of them are bonded to these walls except, of course, for the pier to the southeast behind which there is no longer any original brick wall, but, as was stated before, a tufa masonry of the XV century.[88] By comparing the shape of the piers found above the vault with the present crossing piers, it results that the original piers between nave and transept must have been cross-shaped. Yet it is important to note that the piers in the transept do *not* continue the direction of the clerestorey walls of the nave. They stand a little further outside so that the distance from the N. pier to the S. pier is 60 cm (2 R. ft.) longer than the inner width of the nave. Thus it appears that these piers can hardly have had the structural purpose of abutting the walls of the nave. Of course, they may have carried two of the cross beams of the transept, but certainly such piers would hardly be needed for supporting the beams which, throughout the edifice, are carried by the walls themselves. On the other hand, it is manifest that they definitely subdivided the transept into three bays, one in the center and one in each of the two wings.

It would seem, however, that this arrangement of piers, which run straight up to the top of the walls and end there without any structural purpose, was not what was originally intended. The southeast pier shows, 0.70 m below its top and 14.80 m above the present level of the church, 15.50 m above the original level, four bricks which clearly form the beginning of an arch (Figs. 6, 13). Even the setback, which generally marks the springing of an arch in Roman and Early Christian structures, is clearly visible. Evidently the arch was never completed and the plan to build it was abandoned in favor of the present arrangement with straight piers. Neither the southwest pier nor the two piers to the north show any signs of similar arches, and at the southeast pier itself the pier was continued above these four bricks with straight masonry. Yet there can be no doubt that these bricks indicate the beginning of a voussoir which is bonding into the original masonry. Thus they testify to the existence of an earlier project which would have subdivided the transept more effectively than was later done by the plain projecting piers.

---

[88] See above, p. 375.

Two main solutions are possible in reconstructing the earlier project to subdivide the transept. There is first the possibility that one large arch spanned the transept on either side from east to west. Since the distance between the southwest and the southeast piers is 8.62 m, the radius of such an arch would have been 4.31 m. If the length of the bricks of the arch with 0.45 m is added, the top of the extrados of such an arch would reach a height of about 19.55 m from the present level.[89] The wall above it, even if assumed to be not more than 30 cm high, would reach a level of about 19.85 m from the present pavement and and 20.55, from the original. The height of a corresponding arch to the north cannot be determined exactly, since the length to which the northeast pier protruded and thus the exact dis-

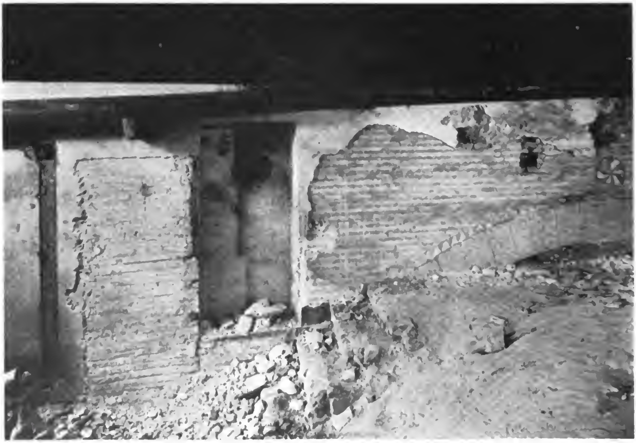


FIG. 13. S. Pietro in Vincoli, transept above vaults, springing of arch between South wing and center bay.

---

[89] 14.80 m (springing of arch) plus 4.31 m (radius) plus 0.45 m (voussoir) = 19.56 m.

tance between the northeast and northwest piers are unknown. Yet since the northeast pier could not have protruded beyond the depth of the XV century pier below, within which the original pier must be contained, the distance was at least 9.87 m and the radius 4.93 m. Consequently the level reached would be about 20.50 m;[90] figuring from the original level of the church, the height reached would be about 21.20 m.[91] That would mean that the walls of the transept which, before the XV century, were 15.25 m high,[92] would have required an additional height of 5 m to conform to this project. This in itself is not very likely. Moreover, in Roman and Early Christian architecture, the voussoir of an arch of such considerable strength would always be composed of two or three rows, not of one single row of bricks, as is proved by the triumphal arch in any Early Christian basilica. And, last but not least, the few bricks that are preserved are laid out so as to indicate clearly an arch of considerably less radius. In other words, it seems certain that the first project did not require one arch but a series of arches between the center bay and the wings of the transept. Of course, such a series of arches would not have rested on free standing piers which would have to have been 15.50 m high and, therefore, much too tall for free standing supports. They could only have been openings in the upper part of a wall which separated the center bay of the transept from the wings; the lower part of such a wall would have rested on an arcade consisting of a number of arches (Fig. 14, top).

The number of these upper openings can be tentatively determined from the inclination of the bricks preserved from the projected arch. Of course, such a calculation which can be based only on such a small remnant cannot be entirely accurate. Yet it seems clear that the fourth brick, which starts 34.2 cm above the springing of the arch, has an inclination of 12½ degrees; the third one, 16.6 cm above the springing, has one of 9½ degrees. Both these angles would lead to the assumption that the arch was planned to have a radius of about 1.15 m,[93] that is, a diameter of 2.30 m. Since the voussoirs are each 45 cm long,

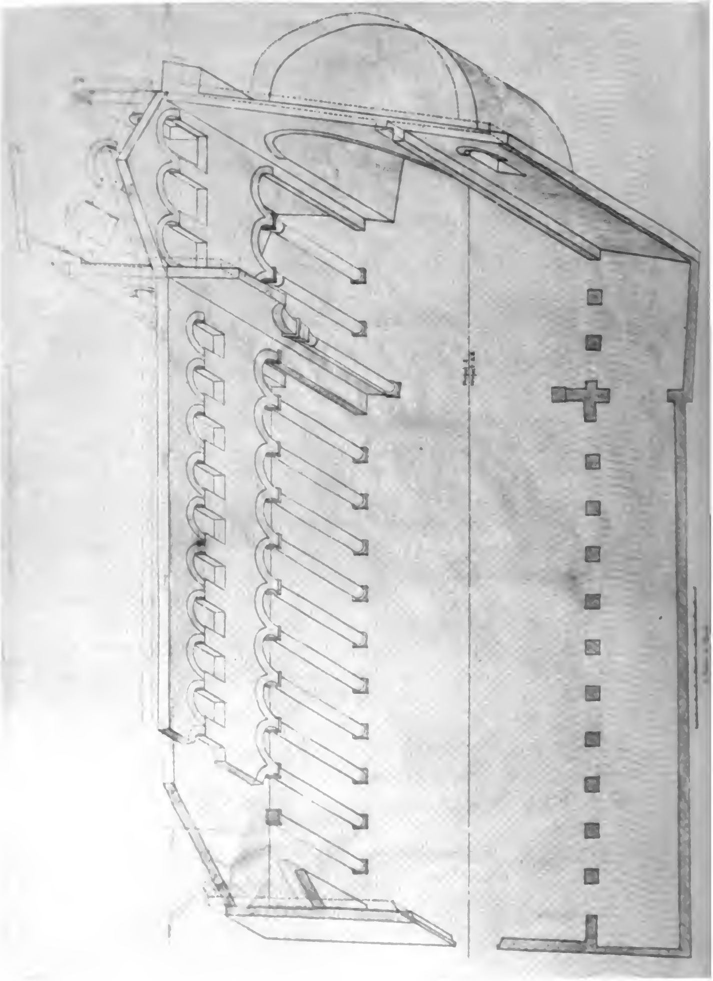
---

[90] 14.80 m (springing of arch) plus 4.93 m (radius) plus 0.45 m (voussoir) plus about 0.30 m (top of wall above voussoir) = 20.48 m.

[91] See above, p. 371.

[92] See above, p. 364.

[93] The exact radius seems to have been 1.17 m.



each arch with its voussoirs would span 3.20 m (11 R. ft.). Since, on the other hand, the distance between the extrados of the two arches east and west on the S. side is 9.52 m,[94] three arches, of approximately 3.20 m each, should have opened in the upper part of the wall which separated the center bay from the S. bay of the transept. Consequently, one will have to assume that likewise three arches would have formed the arcade which supported this wall in the lower part of the transept. The arches in the corresponding wall to the north would have been somewhat wider, the minimum distance between the sustaining piers east and west being 9.87 m and consequently 10.77 m from the extrados of the E. arch to the W. arch. Each arch, including its voussoirs, would span 3.56 m (12 R. ft.). The same width would apply to the arches in the lower arcade, measured from axis to axis of their supports. What these supports would have been like cannot be determined; they may have been either piers or columns. Indeed, the abaci of the columns in the nave are precisely 1.02 m long, and, if similar columns had been intended for use in the transept, they would have been large enough to support two voussoirs of 0.45 m each.

Such an arrangement of three arches in the upper part of the walls between the wings and the center bay of the transept would necessarily lead to the assumption that, in this first project, the outer walls of the transept were higher than they are now. But they need not have been much higher, not more than 1.45 m, so that from the present level of the pavement they would have reached a height of about 16.90 m.[95] The distance from the original level, which was 70 cm lower, would make these walls exactly 17.60 m high (60 R. ft.) (Fig. 14, top, dotted line).

As a matter of fact it seems as if traces of such a higher transept are still preserved in the remnants of a wall which appear to the right above the main apse and which seem to indicate that originally a higher apse either existed or was planned. But in order to clarify this point, the apses as a whole have to be investigated (Fig. 3).

That the lateral apses were added to the original building in the XV century has been discussed before, and it has likewise been stated that, at least in the S. apse, the foundation walls

---

[94] 8.62 m (distance between piers) plus 0.90 m (two voussoirs at 0.45 m. each).

[95] 14.80 m (springing of arch) plus 1.78 m (radius and voussoir) plus 0.30 m (wall top).

are of the XV century. Nor could there have been any earlier lateral apse in place of the N. apse, for the arch above this lateral apse to the north indicates that there was in the original building a window in this place, and the existence of such a window precludes, of course, the simultaneous existence of an apse at the same point.[96] There were definitely no lateral apses in the original edifice. The main apse contains many parts which belong to the original structure, but it does not by any means belong to it in its entirety. It has been frequently restored from the XV century on,[97] but quite aside from these later restorations, different building periods manifest themselves quite clearly (Fig. 6). The top of the wall to about 1.50 m below the cornice shows characteristic High Medieval features. The cornice has a frieze of modillions set between two dentil friezes; the small window to the right is covered with a segmental arch; the masonry has a ratio of 9 bricks to 8 mortarbeds, and the mortar is quite smooth and smeared over the edges of the bricks. All this points to a date in the XII or XIII centuries. One may compare the cornice with that of S. Stefano del Cacco in Rome, the masonry with that of S. Salvatore in Onda which were constructed in the XII century. Below this masonry there appears another which extends down to an irregular line corresponding sometimes approximately to the level of the architraves of the rectangular XVI century windows, while in some places it extends farther down to about half their height. Like the upper masonry, it consists of small bricks of different colors, from a brownish red to a light yellow, and while it is certainly earlier than the masonry on top of it, it is nevertheless not much earlier. It may date from the XI or the early XII century. In this zone traces of a number of small windows are preserved, a first one at the extreme right which had to be walled up when the XV century window was built next to it;[98] one just above the strong buttress in the center of the apse; a third one to the right of the rectangular XVI century window on the S. side of the apse; and a fourth one to the left of this window. One would suppose that originally a fifth window existed just above the rectangular window to the north, where there is now a large section of XIX century masonry. All these small windows are covered with semicircular arches which spring from the jambs without any traces of setbacks. The bricks of their voussoirs are 22 cm long, their openings are approximately 60 cm wide, and, including their arches, about 1.15 m high. Similar windows are frequently found in the early XII century, for example, in the Upper Church of S. Clemente of about 1128 and at S. Stefano Rotondo where they belong to the restorations of 1130–43.

---

[96] See above, p. 376.

[97] See above, p. 357 f.

[98] See above, p. 363.

While both these upper masonries are certainly Medieval, another very different one stretches below them down to about the level of the sills of the XVI century rectangular windows (zone II). Wherever one of the small Romanesque windows had been arranged, this masonry was broken out; the jagged outlines in the brick masonry clearly mark these repairs. The bricks used in this zone are arranged in far more regular layers than in the Medieval parts of the apse, with only a very few yellow bricks interspersed among the red ones. 8 bricks and 7 mortarbeds alternate every 50 cm; the bricks are from 3.7 cm to 4.0 cm high, the mortarbeds from 2.2 cm to 2.9 cm thick. It is a masonry very similar, though possibly somewhat later than that of the transept. The same brickwork was used to repair a large square buttress at the N. corner of the apse.

Yet just this repair of the buttress shows that not even this masonry is the one which forms the original parts of the apse. The top, as well as the bottom part of the buttress, shows a brickwork which, though similar, is even more regular than the one used in zone II. There prevails in it a ratio of 8 or 7½ bricks and 7½ and 7 mortarbeds every 50 cm; the bricks average from 3.0 to 3.8 cm, the mortarbeds from 2.1 to 3.5 cm. The lower part of the buttress is bonded to a narrow vertical strip of the same masonry, which runs along the N. corner of the apse and to a large zone of brickwork of the same type around the foot of the apse. These are the only remnants of the original masonry of the apse proper and they are done in a brickwork which seems to be identical with that of the transept. There are no traces of windows preserved, neither in this original zone of the apse nor in the zone which was subsequently erected on top of it.

This same original masonry continues above the top of the buttress as a broad projecting mass which was cut through at some time. It is about 2.50 m wide and at its right falls exactly in line with the N. flank of the tall buttress at the corner of the apse, while at its left it corresponds precisely to the inner wall of the apse. It reaches a level which almost corresponds to the level reached by the four protruding pilasters inside the transept, and clearly bonds with the wall of the N. transept on either side (Figs. 3, 6).

This projecting mass of masonry can hardly be anything but the remnants of an apse wall and of a buttress which were intended to be 1.50 m higher than the present one. Such an apse may never have been completed, but it certainly existed as a project. If it is recalled that in the transept also, traces of a first unexecuted project are preserved in which the wings were separated from the center bay by walls with arcades and with openings in the upper part, and that also this project would have required transept walls about 1.50 m higher than the ones executed, it becomes very likely that this first plan for the transept and the first plan of the apse belong to one and the same project.

This first plan was evidently given up at a very early period for reasons unknown. Possibly the main part of the apse and part of the corner buttress collapsed and were repaired with the masonry of zone II, since this cannot be dated much later than the original masonry. Simultaneously both apse and transept may have been lowered. In the interior the first project, with arcades and upper openings between the center bay and the wings of the transept, was given up in favor of a second solution which retained only four piers at the corners of the center bay (Fig. 14, bottom). Whether in this second phase an arcade was still arranged or planned between the lower parts of these piers cannot be determined, but it is possible. In the upper parts of the transept, at any rate, the idea of a division of the transept by means of walls was discarded. Simultaneously with this change in plan, the whole transept was completed and the clerestorey walls of the nave may have been begun. It would seem, as a matter of fact, as if the building of the whole church had proceeded from east to west; the fact that the width of the windows in the nave and their distances apart continuously increase from the transept towards the façade seems likewise to point to such a process.

To sum up: The aspect of the edifice in this second plan was preserved up to the XV century with only slight changes. The nave was bounded by eleven arcades resting on Doric columns, set, in all likelihood, on bases. The level of the nave, as indeed of the whole church, was about 70 cm lower than the present one. The clerestory of the nave was lighted by eleven windows, one over each arch. The nave was covered by either an open timber roof or by a flat ceiling. The aisles had also either open timber roofs or flat ceilings; whether their walls had any windows cannot be ascertained, although the analogy with other churches of the period makes it most unlikely. Whether an atrium existed in front of the basilica cannot be determined either, although the existence of the large plaza in front of the present building might support such an hypothesis. While the entrance wall of the church had no doors leading into the aisles, either one or three doors led into the nave. The transept, while orthogonal to the axis of the nave in its western parts, changes its angle in its eastern part. It is clearly bonded to the nave and aisles. Its center part was separated from the wings by four projecting piers which still reflected an earlier project in which wings and center bay would have been separated by arcades and upper walls. The existence of these projecting pilasters in the transept makes it clear that the piers between the nave and transept were always cross-shaped. The center bay of the transept was about 60 cm wider than the nave. The transept wings were lighted by a number of windows, one over the roof of the aisle, two in the N. wall, and two, one above the other, in the E. wall of the N. wing. The arrangement of windows in the S. wing can only be assumed to have been identical. There existed no lateral apses; only one main apse, possibly without windows, opened from the main part of the transept. It was flanked at its N. corner by a strong buttress; a similar buttress may have existed in the S. corner (Fig. 14).

III

The manifest disaccord between the axes of nave and transept of S. Pietro in Vincoli may possibly have been influenced by the existence of earlier edifices on the site of the present church. Indeed, remnants of a number of earlier structures have been found, partly within the walls of the apse where they are still visible, partly below the pavement of the apse and below the center part of the transept. These latter were uncovered when the new confessio was erected in 1876–77; they are no longer visible and the descriptions which are preserved of them are exceedingly poor. De Rossi,[99] who at the time was absent from Rome and thus did not see the excavations, was able to quote only two short reports. The first, made by Father Tongiorgi, describes the finding of a IV century sarcophagus which contained the relics of the seven Maccabeans.[100] The sarcophagus was found placed lengthwise below the predella and the first steps of the ancient altar;[101] its interior had been adapted for the relics by subdividing it by marble slabs into seven small compartments. Two lead tablets with inscriptions referring to the relics were found, one in and one near the sarcophagus.[102] The second report used and published by De Rossi was made by Tomassetti. He mentions only very briefly that "at a distance of about 3 m from the wall of the present apse a fine semicircular wall and a straight wall of equal structure were found." He likewise mentions remnants of what "possibly was a hypocaust one meter below the level of the ancient chancel".[103] De Rossi announced he would make a more thorough report, but he never fulfilled his promise and unfortunately also the Notizie degli Scavi[104] refers only to De Rossi's scanty notes. Only Lefort, in one of his archeological letters,[105] added a few observations made by his correspondent who was on the site. The sarcophagus leaned against a brick wall which supported one of the sides of the altar, and some of the bricks of this wall bore a round stamp, 0.015 m in diameter.[106] Lefort's correspondent read it as "opus DOL.EX. FIG linis PontICULAnis DomIN Nostri" and dated it in the period of Faustina and Marcus Aurelius which would correspond to the years 161–175. This date is certainly somewhat early; the stamp must be dated late II or early III century.[107] Marucchi, who may have seen the excavation when a young man, adds one more detail: the small apse which was found showed remnants of a painted decoration "similar to that of S. Saba".[108]

---

[99] G. B. De Rossi, "Scoperta d'un sarcofago colle reliquie dei Maccabei" in: *Bullettino di Archeologia Cristiana*, ser. III, anno I (1876), p. 73 ff. (quoted as "*B. A. C.*, 1876").

[100] It is now preserved in the confessio below the modern High Altar.

[101] *B. A. C.*, 1876, p. 73 ff.: "Il sarcofago era collocato trasversalmente sotto la predella e i gradini dell' altare isolato di modo che con una delle estremità toccava la base della mensa e con l'altra terminava sotto l'ultimo gradino. . .".

[102] While neither De Rossi nor Tongiorgi attempted to date these inscriptions, O. Marrucchi, *Elements d'Archéologie Chrétienne*, III, Rome, 1902, p. 318, attributed them to the X or perhaps even the IX century.

[103] *B. A. C.*, 1876, p. 73 ff. "Si vede però un bel muro semicircolare tre metri circa distante da quello dell' abside odierna e un muro rettilineo di equale costruzione e residui d'un antico forse ipocausto un metro sotto il livello dell' antico presbiterio."

[104] *Monumenti Antichi . . . dell' Academia dei Lincei*, Notizie degli Scavi, 1876, p. 138.

[105] *Revue Archéologique*, nouv. sér. XXXII (1876), p. 212 f.

[106] *Ibidem*: ". . . le sarcophage joignait un mur très ancien qui supportait l'une des façades de l'autel. . . . En déblayant la crête du mur antique on a arraché deux ou trois vieilles briques portant l'empreinte d'un feu violent et des timbres imprimés sur elles. . . . J'en ai déchiffré un . . . le timbre est inédit. . . .".

Neither of these reports is very helpful in reconstructing the aspect of the structure found, and, as a matter of fact, the strangest possible conclusions have been drawn from them. It is generally assumed that the excavated walls had belonged to an earlier church with its entrance on the side opposite the present one. Marucchi enlarged this statement by tentatively dating this "earlier church" into the IV century,[109] and Lanciani, in his *Forma Urbis,* even drew a small chapel with its apse to the west and its entrance to the east right under the apse of the present church.[110] It is very strange that Lanciani should have made this mistake, for among his own papers there was an exact drawing of these excavations which I was fortunate enough to discover two years ago.[111] This drawing (Fig. 1) carefully marks all the walls found under the pavement. While the stratigraphy of the whole remains unclear, it becomes obvious that not one apse, but at least three, were found and not one straight wall, but several of them. The whole arrangement is very different from what one would expect from reading the reports of 1876, and certainly none of the structures found belongs in any way to any kind of ecclesiastical edifice.

At first glance it seems almost impossible to differentiate the different structures in the drawing. Gradually one realizes that

---

[107] *C. I. L.*, XV, 1, no 405, 5. "Rep. a. 1876 sub altari ecclesiae s. Petri in vinculis . . . aetatis Severianae." The reading is here corrected into "DOMIN. NOSTRORUM".

[108] Marucchi, *op. cit.*, p. 315.

[109] *Ibidem*, p. 315.

[110] Lanciani, *op. cit.*, Pl. 22, 23.

[111] The drawing is preserved in the R. Istituto di Storia dell'Arte e di Archeologia, Palazzo Venezia, Rome, Lascito Lanciani, XXXIX, Vol. I, f. 49, no. 30930. The drawings, nos. 30924 and 30923 (*ibidem*, f. 46 and 47), are, the first one a section, the second one an elevation of the new confessio of S. Pietro in Vincoli. The drawings were evidently done by Count Vespignani, who was the architect of the new confessio; his drawing style, as well as his handwriting, is quite unmistakable. I am very much indebted to the directors of the Institute for giving me access to these drawings. The drawing no. 30930 is incorporated into our plan, Fig. 1. We have added only the hatching and the letters marking the different walls.

there are evidently three different systems to be distinguished which were superimposed at different times. First, there is a strong straight wall, A 1, about 0.95 m thick, which runs in a slightly oblique direction across the apse. It evidently forms one of the main walls of the first system and certainly continued north and south beyond the small area excavated and, in all likelihood, beyond the apse. It is this wall against which the sarcophagus of the seven Maccabeans was leaning (it is clearly marked on the drawing) and which consequently contained the brick with the Severian stamp. Two travertine blocks [112] were set on top of this wall at a distance of 3.60 m from each other. Strangely enough, their faces seem to be set at a right angle to the wall on which they stand. A curvilinear wall, A 2, only 0.50 m thick, forms the beginning of a flat segment shaped apse which was evidently protruding from this wall towards the southeast. On exactly the same axial system with these two walls lies a weaker wall to the south, A 3, which projects in the direction of the nave. Two other walls, A 4 and A 5, 2.75 m west of the main wall, A 1, stand at a right angle to A 3 and parallel to A 1; 3.65 m northwest of A 1 are marked two other walls, A 6 and A 7. While only these walls are marked on the drawing, one more wall of the same network, A 8, is still visible enclosed in the foundation walls of the present apse of the church not far from its S. corner (Fig. 1). Its brickwork shows 10 bricks and 10 mortarbeds every 50 cm with bricks 3.1 to 3.4 cm high and with mortarbeds 1.7 to 2.2 cm high. This masonry resembles other structures of the early III century, such as the original edifice of S. Croce in Gerusalemme,[113] or the Thermæ of Caracalla. Thus it confirms the date of about 200 A.D. suggested by the brickstamp which was found in the main wall, A 1, in 1876.

Somewhat later than this system A is a single wall, A′, which is enclosed in the wall of the present apse at its S. corner. It starts from a level somewhat higher than the level of A 6. Its direction differs slightly from that of system A, and its brickwork with 9½ bricks to 9 mortarbeds in every 50 cm indicates a date about the middle of the III century.

While this wall is evidently an addition to system A, an en-

---

[112] They are marked "Travertino" in the drawing.

[113] R. Krautheimer, Corpus Basilicarum Christianorum Romæ, I Vatican City, 1937, p. 172.

tirely different second system, B, was superimposed on the first one at a later period. That it is later becomes quite evident from the drawing where one sees that the walls of system A were frequently broken out when meeting the structures of system B, for example, the wall of the segment shaped apse, A 2. In other places the walls of the second system B are so weak that they could not have been preserved at the time of the excavation if they had been situated below the strong walls of system A. The preservation of the hypocaust piers to the south, for example, can be explained only if it is assumed that they stand on the main wall of system A; if they had been underneath it, they would have been completely crushed. This system B consists first of the remnants of a small stilted semicircular apse, B 1, to the south, with its opening turned east-southeast; the apse as well as a neighboring room, B 2, are filled with hypocausts. This second room B 2 shows also what appears to be the beginning of an apse. These two rooms are joined to a wall, B 3, east of the main wall, A 1, of the first system, and this wall in turn leads to a long wall, B 4. B 4 runs due west-northwest and is connected with a third semicircular apse, B 5, again with remnants of hypocausts. These latter walls, B 3, 4, 5, are laid out on an axial system, which differs slightly from that of the first two rooms, B 1 and B 2; yet the difference is so slight that they seem to belong to the same network. According to Vespignani's notes on his drawing, the pavement of this whole system was situated 2.15 m below the level of the ancient chancel, which corresponds to a level of 1.80 m below the pavement of the present nave. While these walls are known only through the drawing, the remnants of two other walls of system B are still visible. Both are preserved on the outside of the present main apse of the church, one, B 6, south of the strong low buttress in its center, the other one, B 7, north of it. While the first one is preserved to a height of almost 4 m, the second one is preserved only in the foundation wall proper of the present apse (Figs. 1, 6). Their brickword seems identical; 8½ bricks and 8 mortarbeds alternate in every 50 cm, the bricks are 3.4 to 4.3 cm high, the mortarbeds 1.7 to 2.1 cm thick. This brickwork also seems to point to a date later than that of the brickwork of system A. Similar brickwork occurs in the Aurelian walls of about 275,[114] and as late as the early IV century, when the first hall of S. Crisogono was built.[115]

While it is obvious that neither of the two systems, A and B, is in any way connected with the structure of the present church, it seems impossible to attempt a reconstruction of the two edifices to which these remnants belong. The structures of system B were in all likelihood part of some private thermal edifice, as would appear from the frequent use of apses and from the use of hypocausts. Yet this edifice was not connected in any way with either the Thermæ of Trajan or the Thermæ of Titus which extend one east, the other one south on the hill on which S. Pietro in Vincoli is situated. Both run on different axes and both are too far distant. The scale of the building B seems quite small and would suggest rather some private thermæ. As regards the first edifice, A, it is even harder to be definite. It also may have belonged to some thermæ, but it may as well have formed part of a residence; segment shaped apses would be possible in either type of edifice. The street pavements found in the neighborhood of the church unfortunately are insufficient to give further evidence.[116]

In contrast to these two networks, A and B, a third group of structures, marked on Vespignani's drawing, seems to conform definitely to the axis of the nave of the basilica. Only two walls are preserved of this system, one, C 1, which runs east to west along the S. side of the present confessio, and the other, C 2, which departs from C 1 at about its middle at a right angle and runs due south. They are certainly superimposed on the structures of both systems A and B since they overlap these earlier walls several times. It seems that they are foundation walls for choir screens which would have consisted of a center part projecting from the apse into the transept (C 1 would have formed the S. wall of this center part) and of lateral screens which did not project so far and crossed the transept transversally. The transverse screens would have been similar though earlier than those visible on Rossini's engraving[117] and on the groundplans done before 1876.[118] The date of the excavated screen foundations is uncertain, but since their axis corresponds to that of the nave, they must have been erected when or after the church had been completed. On the other hand, they were destroyed when a large Medieval tomb D was placed in the transept.[119] The shape of these screens, with projecting center part, would have corresponded exactly to that of the choir screens at S. Stefano in Via Latina, a church which was constructed between 440 and 461.

---

[114] 8½ bricks to 8 mortarbeds; bricks 3.3 to 3.6 cm high, mortarbeds 2.1 to 3 cm thick.

[115] Krautheimer, *op. cit.*, p. 147.

[116] Lanciani, *op. cit.*, Pl. 22 and 23. Seroux d'Agincourt, *op. cit.*, I, p. 35, mentioned that he had excavated some mosaic pavements and some walls behind the apse. He believed them to have belonged to the Thermæ of Trajan which is obviously erroneous; they may have belonged to either of the earlier edifices buried below the apse.

When the church was built, walls A 8 and B 6 must have been still preserved to a considerable height and their remnants incorporated into the wall of the apse. Yet it seems even more important that, in building the apse and the transept, the strong main wall, A 1, was evidently dug up and, in its continuation outside the apse, was used as the base for the E. wall of the S. transept and the large buttress at the N. corner of the apse. This explains the oblique direction of the E. wall of the transept and the oblique direction of the opening of the apse. Consequently also, the axis of the apse had to be fitted into this system. It is possible that the direction of the N. wall of the transept, although it does not quite fit into this same network A, was also determined not only by the direction of an ancient street corresponding to the Via delle Sette Sale but also by some ancient wall which has not yet been discovered. Only through their relation to some underlying walls can the direction of the transept walls be explained. Not before the W. wall of the transept was built was this older system discarded and replaced by a new one to which the nave and aisles were subordinated.

IV

The question of the date of the building of S. Pietro in Vincoli remains to be discussed. To determine it, different indications might be considered. Aside from the documentary sources, the brickwork and proportions of the windows give very definite clues. Since the bricks are frequently pilfered from older constructions, it seems safer to rely on the character of

---

[117] See above, p. 358 and Fig. 4.

[118] See above, n. 9.

[119] Marked on Vespignani's drawings "Sepultura medieovale". There is no way of dating it precisely.

BRICKWORK

Date	Church	Height of Bricks	Average	Height of Mortar	Average	Ratio
368–99	S. Clemente	2.5–4.4 cm	3.65 cm	1.3–5.6 cm	3.2 cm	8/7½ : 7½/7
350–400	S. Croce (apse below)	2.8–4.3 cm	3.5 cm	2.5–4.7 cm	3.6 cm	8/7 : 8/7
418–32	S. Sabina	3.5–4.0 cm	3.7 cm	3.0–5.1 cm	4.0 cm	7½/7 : 7/6
432–40 or slightly earlier	S. Maria Maggiore	3.5–4.5 cm	3.9 cm	3.5–6.0 cm	4.7 cm	7 : 6
432–40	Lateran Baptistery	3.0–3.5 cm	3.3 cm	2.4–3.7 cm*	3.1 cm	8 : 7
432–40	S. Lorenzo in Lucina	2.3–3.8 cm	3.1 cm	2.0–4.8 cm	3.3 cm	8 : 8
410–61	S. Giovanni e Paolo	3.5–4.2 cm	3.7 cm	2 –4.2 cm	3.1 cm	7½/7 : 7
	S. PIETRO IN VINCOLI transept	3.5–4.5 cm	3.7 cm	2.2–4.0 cm	3.3 cm	8 : 7½/7
	apse, zone I	3.0–3.8 cm	3.4 cm	2.1–3.5 cm	2.7 cm	8/7½ : 7½/7
	apse, zone II	3.7–4.0 cm	3.8 cm	2.2–2.9 cm	2.6 cm	8 : 7
476 or earlier	S. Agata dei Goti	2.8–4.8 cm	3.4 cm	1.8–3.5 cm	2.9 cm	8 : 8
468–83	S. Stefano Rotondo	2.7–5.1 cm	3.7 cm			8 : 7

* My measurements differ slightly from those given by G. B. Giovenale, *Il Battistero Lateranense*, Rome, 1928, p. 43.

WINDOWS

Date	Church	Width	Distance	Cumulative Distance in R. Ft.	Ratio
401–17	S. Vitale	2.10 m	0.90 m	10 R. ft.	1 : 0.43
418–32	S. Sabina	2.43 m	1.18 m	12 R. ft.	1 : 0.49
432–40 or slightly earlier	S. Maria Maggiore	2.00–2.46 m	0.89–1.19 m	10–12½ R. ft.	1 : 0.49–1 : 0.50
430–40	S. Giovanni e Paolo (S. wall)	1.75 m	1.25 m	10 R. ft.	1 : 0.7
440–61 or slightly earlier	S. Giovanni e Paolo (N. wall)	1.55 m	1.43 m	10 R. ft.	1 : 0.92
	S. PIETRO IN VINCOLI (E. bays)	1.78 m	1.38 m	10¾ R. ft.	1 : 0.77
	S. PIETRO IN VINCOLI (W. bays)	2.02 m	1.50 m	12 R. ft.	1 : 0.74
476 or earlier	S. Agata dei Goti	1.70 m	1.35 m	10 *Byz.* ft.	1 : 0.8
468–83	S. Stefano Rotondo	1.60 m	1.90 m	12 R. ft.	1 : 1.2

the mortarbeds, on the thickness of the mortar used, and on the ratio between bricks and mortarbeds within a given height. In the original parts of the apse and of the transept of S. Pietro in Vincoli, these elements point to a date somewhere in the first half or about the middle of the V century. The ratio of between 8 and 7½ bricks and between 8 and 7 mortarbeds every 50 cm, the height of the bricks, which fluctuates from 3.0 to 4.5 cm, and the thickness of the mortarbeds, which varies between 2.1 and 4.0 cm with an average thickness of 3.3 cm, corresponds to the brickwork of Roman structures which are dated between 400 and 450. A list of brickwork measurements is added from which it would appear that the masonry of S. Pietro in Vincoli is most closely related to that of S. Lorenzo in Lucina [120] and that of Lateran Baptistery,[121] both erected or consecrated between 432 and 440. Yet since there is no building preserved in Rome which would show the development between 450 and 470, a margin for error has to be left. Indeed the proportions of the windows of S. Pietro in Vincoli point rather to a date about the middle of the V century, for it seems that in the Roman churches of the V century, the ratio between the width of windows and their distances from each other undergoes a very characteristic change about the middle of the century. During the first forty years of the century, the proportion sinks from 1:0.43 at S. Vitale (401–17) to 1:0.50 at S. Maria Maggiore (before 432), that is, the window is first more than twice as wide and then just twice as wide as the distance from one window to the other. About 440, in the S. clerestorey of SS. Giovanni e Paolo, the ratio sinks to 1:0.7, that is, the distance between the windows increases. This development continues throughout the second half of the century until the window is smaller than the pier. Again the accompanying list will demonstrate this development and the place which S. Pietro in Vincoli occupies within it.

On the other hand, the tradition for the foundation of the church seems to be well supported by documentary sources. They have been carefully collected by De Rossi [122] and by

---

[120] R. Krautheimer and W. Frankl, "Recent Discoveries in Churches in Rome," in: *American Journal of Archaeology,* XLIII (1939), p. 388 ff.

[121] G. B. Giovenale, *op. cit.,* p. 43.

[122] G. B. De Rossi, *Inscriptiones Christianae Urbis Romae,* II, 1, Rome, 1857–88 (quoted as "De Rossi, *Inscriptiones*").

Kirsch.[123] The interpretations given to these sources, however, may sometimes be open to revision.

What we know for certain out of a mass of legends and traditions are a few main facts:

A. In 431, Philippus, a Roman presbyter and the leader of the Roman delegation to Ephesos, signs his name under the resolutions of the Council which he had attended and in which he had played a decisive role. According to this signature he was "presbyter ecclesiae apostolorum", presbyter of the Congregation of the Apostles.[124]

B. A whole set of inscriptions originally decorated the interior of S. Pietro in Vincoli. They are all lost, but they are preserved in different collections of inscriptions ranging from the VII to the XV century. They are:

a. An inscription which is preserved in the Sylloge IV of Lorsch of the VII century and in the Sylloge of Verdun of the VIII century, in the latter with the remark that it was to be read "in occidentali parte ipsius ecclesiae," in the western part of this church. De Rossi has suggested that this refers to a place above the door on the interior façade.[125] The text of this inscription was:

Cede prius nomen novitati cede vetustas
regia laetanter vota dicare libet
haec petri paulique simul nunc nomine signo
xystus apostolicae sedis honore fruens
unum quaeso pares unum duo sumite munus
unus honor celebrat quos habet una fides
presbiteri tamen hic labor est et cura philippi
postquam effesi XPS vicit utrique polo
praemia discipulus meruit vincente magistro
hanc palmam fidei rettulit inde senex.
(Syll. Lauresham. IV, f. 71'.)[126]

---

[123] J. P. Kirsch, *Die römischen Titelkirchen im Altertum,* Paderborn, 1918, p. 45 ff.

[124] J. Hardouin, *Conciliorum collectio regia maxima,* I, Paris, 1715, col. 1483 f.

[125] De Rossi, *Inscriptiones,* II, 1, p. 134, no. 3.

[126] *Ibidem,* p. 110, no. 67; see also E. Diehl, *Inscriptiones Latinae Christianae Veteres,* I, Berlin, 1925–31, p. 182, no. 974. The Syll. Virdun., De Rossi, *Inscriptiones,* II, 1, p. 134, gives a text with slight variations. As has been pointed out by De Rossi, ''Epigrafe d'una chiesa dedicata agli apostoli Pietro e Paolo,'' in: *B. A. C.,* ser. III, anno 3 (1878), p. 14 ff., and by St. Gsell, *Les Monuments antiques de l'Algérie,* II, Paris, 1901, p. 159 f., this inscription was copied in the VI century in a church at Ain-Ghora near Tebessa, in Africa, with alterations adapting it to the local circumstances.

This inscription reveals a number of facts: The church had formerly a different name and was rebuilt (line 1), evidently with the support of the Imperial house (line 2); it was dedicated to Saint Peter and Paul (line 3) by Sixtus (which must mean Sixtus III (432–40) who was a contemporary of the Council of Ephesos) (line 4). Particular emphasis is given to the equality of the two Apostles and to the unity of creed (lines 5 and 6). Philippus, who is responsible for the building and in charge of the church (line 7), brought back, when an old man, the palm from Ephesos where Christ, his master, had been victorious (line 8).

It is obvious from this inscription that Philippus, who was the pontifical legate at Ephesos and who signed his name there as presbyter of the Congregation of the Apostles, was the same priest who is mentioned in this inscription.[127] It is likewise obvious that the present church of S. Pietro in Vincoli is either identical with or the legal successor of the "ecclesia apostolorum" mentioned in the Ephesos document. It is further likely that an older ecclesiastical building must have existed on the site of S. Pietro in Vincoli, as proved by line 1, and that between 432 and 440 a dedication of a new church took place. This dedication was made to Saint Peter and Saint Paul and replaced an older one. Yet a number of questions remain unanswered.

First: which name did the older building have before it was replaced by the new church, and which exactly was the new title given the new church by Sixtus III? The general assumption is that, before Sixtus III, the older building had been dedicated to the twelve Apostles, that its name had been Ecclesia Apostolorum, the name used by Philippus at Ephesos in 431, and that the new name given it by Sixtus was Saint Peter and Paul.[128] I wonder whether this assumption is really conclusive. Of course, in 431 at Ephesos, the church is called ecclesia apostolorum, but this same name is used only once, namely in this very document, before the new consecration took place—and indeed very shortly before this consecration—while it occurs frequently after the consecration, from the V through the VIII century and up to the IX. One might try to explain this by assuming that the new name of Saint Peter and Paul was not universally accepted and that the old name of the Apostles persisted. Yet that cannot be maintained because it is only in official documents that the dedication to the Apostles continued, while the unofficial documents used entirely different names.[129] In all legal documents, for example, in the signatures of the Roman synods of 499[130] and of 595,[131] the church remained the "titulus Apostolorum" or "titulus Sanctorum Apostolorum" and this very name occurs frequently in the official biographies of the popes in the Liber Pontificalis.[132] Not once, on the other hand, in any document is the church called St. Peter and Paul. There is only one explanation. Saint Peter and Paul are the "principes Apostolorum"; they are, in the eyes of the Early Christians, the Apostles *kat' exochēn;* their memoria at S. Sebastiano f.l.m. in Rome is known since the IV century as Memoria Apostolorum;[133] the church which Rufinus, the powerful minister of Theodosius I, founded about 393 at Chalkedon near Constantinople and which was consecrated to Saint Peter and Paul was called the Apostoleion;[134] the church of the Apostles in Paris, built in 512, was likewise dedicated to Saint Peter and Paul.[135] It would be easy to increase the number of instances. Thus it seems safe to assume that also at S. Pietro in Vincoli the name of Saint Peter and Paul and the name of the Apostles were synonymous, and that, consequently, the name "ecclesia apostolorum," as used in 431 by Philippus at Ephesos, was identical with the name of Saint Peter and Paul as used in the inscription of Sixtus III between 432 and 440. The continuous use of the term "titulus Apostolorum" in legal documents referring to the church up to the IX century, the fact that none of these documents uses the name of Saint Peter and Paul; and finally the constant synonymity of the names of Saint Peter and Paul and of the Apostles throughout the Early Christian period make such a solution imperative.

---

[127] For his biography see also A. F. von Pauly and G. Wissowa, *Real-Encyclopädie der Klassischen Altertumswissenschaft*, XIX, 2, Stuttgart, 1894 ff., col. 2375 f.

[128] De Rossi, *Inscriptiones*, II, 1, p. 134; Kirsch, *op. cit.*, passim, and others.

---

[129] See below, p. 407 f.

[130] Monumenta Germaniae Historica (quoted as "M. G. H."), *Auct. Antiqu.*, XII, p. 410 ff.

[131] M. G. H., *Epist.* I, p. 366 f.

[132] L. M. O. Duchesne, *Le Liber Pontificalis,* Paris, 1886–92, passim (quoted as "*L. P.*").

[133] H. Lietzmann, *Petrus und Paulus in Rom,* 2nd ed., Berlin, 1927, passim.

[134] L. M. O. Duchesne, "Inscription chrétienne de Bithynie," in: *Bull. Corr. Héllén.*, II (1878), p. 289 ff.

[135] J. Hubert, *L'Art Pré-Roman,* Paris, 1938, p. 8 f., particularly p. 9, n. 1.

If this is the case, however, it is of course obvious that the new name, St. Peter and Paul or the Apostles, was not chosen by Sixtus III. Philippus had used it at Ephesos in 431 before Sixtus III had obtained the pontificate. The church, therefore, must have been given its new name prior to that time, that is, at least shortly before its consecration. There is nothing to contradict such an assumption. Consecration and name giving need not take place at the same time, nor does a consecration necessarily mean that the building is either just begun or already completed. It can take place at any time during the process of building, whenever an important part of the building has been completed, or whenever an important occasion for a consecration offers itself, such as the visit of some high ecclesiastical or secular personality.[136] It is perfectly possible that the building of the church had been started before Philippus went to Ephesos. Indeed, it is obvious that Sixtus III, in his inscription, does not claim any important part in the erection of the building; he speaks only of his part in the consecration. He even does it in a way which suggests that he only gave his blessing to a work undertaken by somebody else before he himself had come to power, a fact which is even more significant when one recalls that Sixtus III was not modest in increasing his fame by appropriating for himself buildings begun by his predecessors or by other persons. S. Maria Maggiore and S. Sabina are well-known instances. Certainly Sixtus consecrated the basilica sometime between 432 and 440, possibly in 439 as will be discussed later.[137] The building may very well have been started between 420 and 430. When this consecration took place, part of the building must have been completed, and these parts must have included at least the lower parts of the W. façade where the inscription was placed according to the Sylloge of Verdun. The natural place for such an inscription would be the space between the door zone and the clerestorey zone across the façade, as can be seen at S. Sabina and as it is known to have existed at S. Maria Maggiore. Since the apse and the transept, as we have seen, were probably completed before the nave, they must have been finished also at the time of the consecration. This does not prove anything, however, in regard to the final completion of the nave. The clerestory might have been still unfinished, with the lower parts protected by a temporary roof, which would explain the fact that the proportions of the windows and their distances apart seem to point to a somewhat later date.

---

[136] The best illustration for this procedure is the voyage of Urban II in 1095 and 1096 from Rome to Clermont Ferrand. On his way he consecrates almost every church building which he finds no matter whether it was completed or just begun or whether only its site had been laid out.

[137] See below, p. 402.

The aspect of the building which preceded this basilica remains obscure and so does its name. As to its plan, it may have been anything, a simple hall such as the first building of S. Crisogono,[138] or a community house which had been transformed from either a private residence or from an apartment house, as was the case at S. Clemente[139] or at S. Giovanni e Paolo.[140] The structures found under the apse, particularly in the later system B, may or may not have belonged to a private residence which was later used for such a community house. Nothing definite can be said until further excavations can be made. As to the name, we can for the present suggest only one possibility. Line 5 and line 6 of the inscription very strongly emphasize that the two great Apostles should condescend to *share* the dedication between themselves. It sounds almost like an elaborate excuse to one of them.

b. A second inscription is preserved in one Codex only, in the Sylloge IV of Lorsch of the VII century.

> "Theodosius pater Eudocia cum coniuge votum
> cumque suo supplex Eudoxia nomine solvit"
> (Syll. Lauresham. IV, f. 71)[141]

The place where it was to be read is not mentioned in the Sylloge nor is even the church where it was to be found; but since it immediately precedes the inscription, "Cede prius nomen,"[142] the existence of which at S. Pietro in Vincoli is ascertained, there is no need to doubt that this inscription also belonged to the set of commemorative inscriptions at S. Pietro in Vincoli. The meaning of this inscription seems obvious. Theodosius II, together with his wife, Eudocia, and his daughter, Eudoxia, fulfilled a vow regarding the church, either in his name or in that of his daughter.[143] Indeed, the dates of Theodosius' reign, of Eudocia's life, and of the life of Eudoxia the Younger seem to fit perfectly into the period of Sixtus III, 432–440. Theodosius became Caesar in 418 and died in 450. He never visited Rome. Eudocia was married to him in 421. In 440 she took up her residence in Jerusalem, possibly in disgrace, and died in 460.[144] Eudoxia the Younger,[145] born in 422, was married to Valentinian III in 437 at Constantinople. She came to Rome in 439 and stayed there until 440.[146] She returned in 449 and stayed until 455 when, after the death of Valentinian III and a few months later that of her second husband, Petronius Maximus, she was carried by Geiserich as a prisoner to Africa. She was ransomed in 462 and lived in Constantinople until her death, the date of which is unknown. Her first visit to Rome in 439 and 440 seems to tally with the inscription of Sixtus III, mentioned before, in which he refers to vows made by the Imperial house. Thus it seems safe to assume that these two years are the time in which both the inscription of Sixtus III and that of Theodosius and his family were put up and when the consecration of the church took place.[147] Indeed, it is perfectly plausible to assume that Eudoxia's visit to Rome was the occasion seized upon by Sixtus III for consecrating the basilica in which the Imperial house had shown its interest. The whole tenor of the inscription, together with the fact that Eudoxia's husband, Valentinian III, is not mentioned in it, seems to justify the assumption that the donations to the church had been made by Theodosius, and that Eudoxia represented him at the dedication festivities in her capacity as his daughter.[148]

---

[138] Krautheimer, *Corpus*, p. 144 ff.

[139] E. Junyent, *Il titolo di San Clemente*, Rome, 1932, passim.

[140] *Ibidem*, p. 107 ff.

[141] De Rossi, *Inscriptiones*, II, 1, p. 110, no. 66; Diehl, *op. cit.*, I, p. 348, no. 1779.

[142] See above, p. 397.

---

[143] This is the usual interpretation given, for instance, by De Rossi, *Inscriptiones*, II, 1, p. 110, and *Kirsch, op. cit.*, p. 49.

[144] Pauly and Wissowa, *op. cit.*, VI, col. 906 ff.

[145] *Ibidem*, col. 925 ff.

[146] Diehl, *op. cit.*, I, p. 348, dates this visit in 438, Pauly and Wissowa, *l.c.*, in 439.

[147] Pauly and Wissowa, *l.c.*, assume that the church was founded in this year, a theory which is certainly erroneous.

[148] This whole interpretation is based on the assumption that in the second line of the distich the ''a'' in ''Eudoxia'', in spite of its being short, is meant to be an ablative, as has always been assumed. This would not be a classical usage, but anything is possible in V century poetry. If this assumption, however, should not hold and if the ''Eudoxia'' were meant to be a nominative, as the short ''a'' would indicate in classical Latin, another interpretation would have to be found. In this case one would have to assume that in the first line a ''fecit'' would have to be supplemented meaning: ''Theodosius with his wife Eudocia made a vow,'' while in the second line the ''cum'' would have to be referred to ''suo nomine'' so as to

c. A third inscription also appears in only one Codex, in the Sylloge of Verdun of the VIII century:

> ''In medio regum celestem respice regem
> nec desunt tua signa fides antistite xisto''
> (Syll. Virdun., f. 212, col. 2)[149]

It is certain that this inscription points again to the pontificate of Sixtus III, to the years 432 to 440. It is furthermore clear that it refers to some mosaic or mural with which it was connected.

It is not certain, however, what the mosaic or mural represented[150] and where it was placed. The Sylloge of Verdun, after having quoted an inscription in the main apse,[151] says that this one was to be read ''in altera absida in eadem ecclesia,'' and it has always been taken for granted that this topographical

---

mean ''and Eudoxia with his name (or her name) fulfilled it,'' not a very classical Latin either.

This interpretation would possibly lead to a slightly different dating. If Eudoxia fulfilled a vow made by her father, in his name, this might indicate that Theodosius was already dead. If this were so, the date of the inscription would fall between his death in 450 and Eudoxia's captivity in 455, during Eudoxia's second stay in Rome. This would point to the possibility that work at the church was still going on at this time, an assumption which seems to be supported at least by the measurements of the clerestory.

[149] De Rossi, *Inscriptiones*, II, 1, p. 134, no. 2.

[150] De Rossi, *l.c.*, and Kirsch, *op. cit.*, p. 49, had assumed that the representation showed Christ, the king of Heaven, accompanied by the ''reges'', the imperial family, and by the pope—the ''antistites xistus''—who carried the ''signa,'' the cross. This is hardly the right interpretation. In the first place, ''signa'' can hardly be connected with ''fides'' to mean sign of faith, that is, cross. Cross is always ''signum,'' while the plural ''signa'' can only have the meaning of signs, miracles. On the other hand, ''fides'' is a nominative or a vocative, not a genitive. Thus the translation of this line ought to be ''there are not wanting your signs, Oh Faith, as long as Sixtus is pope.'' Consequently, there is no reason to assume that Sixtus III was represented at all or that he was represented with a cross on his arm. To represent a living person, even a pope, next to Christ, is a motif which to my knowledge does not occur anywhere before the VI century when it appears at SS. Cosma e Damiano (523–529). This reticence in representing living persons next to Christ has to be kept in mind in interpreting the first line. No doubt Christ was represented as the king of Heaven and amidst the ''reges''. But these ''reges'' were hardly the rulers of this world. If they had been, the term used would probably be the one chosen in the contemporaneous inscription of Valentinian III at S. Croce in Gerusalemme, 429–44: ''reges terrae''. ''Reges'', without any specification, would refer instead, I think, to the Four and Twenty elders offering their crowns to Christ seated on his throne. This is indeed a theme which is very frequently represented in the V century, for example, at the triumphal arch of S. Paolo f.l.m., where it is dated in the decade between 440 and 450, through the inscription referring to Galla Placidia (425–50) and Leo the Great (440–61). It would seem that the representation referred to at S. Pietro in Vincoli would have resembled the slightly later one at S. Paolo f.l.m.

[151] See below, p. 405.

indication refers to one of the lateral apses.[152] The structural analysis of the edifice, on the other hand, has shown that there was definitely no lateral apse in the original edifice and none before the XV century in either the N. or in the S. transept. Thus the interpretation of the remark, "in altera absida", becomes rather puzzling. However, it can be explained if it is remembered that the term "absida" in VIII and IX century Latin, which is the period of the sylloge Virdunensis, is used rather indiscriminately. It may refer to lateral rooms of any shape which open onto a main room,[153] to a rectangular chapel,[154] or to a room opening with an arch onto a main room from which it receives its light.[155] If this looseness of the term is kept in mind, it will appear that at S. Pietro in Vincoli the term might have been applied to the transept wings which were smaller rooms separated from the center bay, or to some lateral chapel. It may possibly refer even to the center bay of the transept, which is a room separated from the nave but opening onto it through an arch and which might thus perhaps have been called a second apse. Of course, such a use of the term "absida" seems rather strange, but on the other hand, one can hardly imagine that a mosaic or mural with the important theme of Christ and the Four and Twenty Elders would have been shoved aside into some small chapel or into one of the wings of the transept. Analogies show that this representation usually occupies either the main apse as it did at S. Croce in Gerusalemme, or the triumphal arch as is the case at S. Paolo f.l.m. Thus it seems worth while to consider at least the possibility that the place referred to was somewhere in the center bay of the transept.[156]

[152] De Rossi, *l.c.;* Kirsch, *l.c.*

[153] Walafrid Strabo, *De Rebus Ecclesiasticis,* c. 6: "Exedra est absida quaedam separata modicum quidem a templo vel palatio . . ." (J. von Schlosser, *Schriftquellen zur Geschichte der karolingischen Kunst,* Vienna, 1896, p. 4).

[154] *Translatio S. Filiberti,* c. 29; the crypt of S. Philibert de Grandlieu, which consists of an ambulatory with three rectangular chapels, is described as: "locus sepulturae mirifice . . . transvolutus tribus perinde absidis circumcirca adjectis . . ." (Schlosser, *op. cit.,* p. 215).

[155] Hrabanus Maurus, *De Universo,* lib. XIV, c. 23: "absida . . . eo quod lumine accepto per arcum resplendeat . . ." (J. P. Migne, *Patrologia latina,* vol. CXI, col. 403); the passage is taken from Isidorus of Seville, lib. IV, c, VIII, No. 7 (*Isidori. . . . Etymologiarum . . . libri XX,* ed. W. M. Lindsay, II, Oxford 1911).

[156] The most convincing solution, which unfortunately remains purely hypothetical, would be that "in altera absida" is a lapsus calami of the scribe and that the line should have read "in arcu abside". To check this reading is, at present, quite impossible.

d. The following inscription may possibly be connected with the three V century inscriptions which have been discussed so far. It runs:

> "Inlesas olim servant haec tecta cathenas
> vincla sacrata petri ferrum pretiosius auro."
> (Syll. Virdun., f. 212, col. 2.)[157]

It appears for the first time in the Sylloge of Verdun in the VIII century under the heading "isti versiculi scripti sunt ad sci Petri vincula".[158] In the IX century, it is mentioned in the Sylloge of Würzburg[159] and in Alcuin's collectanea,[160] and later in numerous codices. It was still preserved about 1420 when it was copied by Signorili with the remark that the verses were "scripti de musaicis in tribuna . . . ",[161] and in 1454 when it was copied by Cyriacus of Ancona.[162] Likewise, Petrus Sabinus must have seen it when preparing his collection of inscriptions of Rome. He quoted it and added that it used to be read in the apse, "In apside templi s. petri ad vincula . . . legebantur", that it was done, "ex opere vermiculato", that is, it consisted of mosaic, and that it was written in very old lettering, "vetustissimis litteris". When he actually wrote his book in 1493–94, the inscription had evidently disappeared, as is proved by his use of the past tense.[163]

It has always been assumed that this inscription was contemporaneous with the set discussed so far. This may be so, but it cannot be proved beyond doubt. Relics of chains, which were popularly believed to be those worn by Saint Peter, must have been in the church as early as 501, when, for the first time, the church appears under the name *a vincula sancti Petri Apostoli*,[164] and the inscription in the main apse, where these chains were kept, is undoubtedly of official character. Yet, it may be either contemporaneous with the erection of the church, somewhat later or even considerably later. If the latter were correct, the inscription merely followed the popular belief. The history of the relics gives no additional clue.

---

[157] De Rossi, *Inscriptiones*, II, 1, p. 134, no. 1; Diehl, *op. cit.*, I, p. 348, no. 1781.

[158] See preceding note.

[159] De Rossi, *Inscriptiones*, II, 1, p. 157, no. 10.

[160] *Ibidem*, p. 286, no. 11.

[161] *Ibidem*, p. 134, adnotatio ad no. 1.

[162] *Ibidem*, p. 352, no. 1.

[163] It may have disappeared during the restorations undertaken by Sixtus IV, as suggested by De Rossi, *Inscriptiones*, II, 1, p. 134.

[164] Duchesne, *L. P.*, I, p. 261 and p. 265, n. 13.

It remains uncertain whether the chains were old Roman property [165] or whether they had been brought from Jerusalem, as later legends assume.[166] Likewise it remains uncertain whether from the very outset, they were believed to be the chains of only Saint Peter or of both Saint Peter and Saint Paul. Relics of chains of Saint Peter were venerated in Rome as early as the first half of the V century, as is proved by a relative inscription at Spoleto, set up by bishop Achilles who is known to have lived about 419.[167] Chains of the Apostles are mentioned in 519 when the Ambassador of Justin I asks for relics of them.[168] In 594 Gregory the Great writes to the Empress Constantina that he would be willing to let her have some filings from the chains which Saint Paul had worn around his neck and on his hands, "in collo et in manibus".[169] The place where all these chains were kept is uncertain. Yet it is a curious coincidence that in the XII century the Mirabilia clearly mention chains of both Saint Peter and Saint Paul being kept at S. Pietro in Vincoli; [170] as late as 1420 Signorili believed the chains with their collar, at S. Pietro in Vincoli, to be those of Saint Paul, "cathena cum collaro . . . qua fuit legatus S. Paulus",[171] in spite of the fact that the inscription in the apse, which he quotes, referred to Saint Peter. In none of the early documents, except in the inscription of Spoleto, are the chains of Saint Peter named explicitly. Yet it is obvious that from the end of the V century on, a set of chains, kept at S. Pietro in Vincoli, was at least popularly believed to be connected with the memory of Saint Peter.[172] On the other hand either chains of Saint Paul must have likewise been venerated in the church or else one and the same chains must have been attributed sometimes to Saint Peter and sometimes to Saint Paul. There is one point which seems almost to speak in favor of the hypothesis that the attribution was fluctuating: the fluctuation in the name of the church itself.

---

[165] H. Grisar, "Der mamertinische Kerker . . ." in: *Zeitschrift für Katholische Theologie*, XX (1896), p. 102 ff.

[166] The later legends from the VIII century sermon of Paulus Winfridus (Hom. 38, ed. Migne, *Patrologia latina*, Vol. XCV, col. 1485 ff.) and the XII century Mirabilia Romae (H. Jordan, *Topographie der Stadt Rom im Altertum*, II, Berlin, 1871, p. 623 f.) through Armellini, *op. cit.*, p. 208 f., connected the transfer of these relics from Jerusalem with the name of either Eudoxia the Elder or of Eudocia or of her daughter, Eudoxia the Younger. None of these traditions seems to be correct. Neither Eudoxia the Elder nor Eudoxia the Younger was ever in Jerusalem; Eudocia was there in 438 and again from 440–461, but Socrates, who very carefully reported this pilgrimage, does not mention any relics which she brought from there (Socrates, *Hist. Eccl.*, VII, 47, 3, ed. Migne, *Patrologia graeca*, Vol. LXVII, col. 839 f.). Marcellinus in his Chronicon, ad ann. 439.2 (M. G. H., *Auct. Antiqu.*, XI, p. 80), mentions only relics of Saint Stephen which the Empress brought to Constantinople.

[167] De Rossi, *Inscriptiones*, II, 1, p. 113, no. 79–81 and *B. A. C.* ser. II, anno II (1871), p. 118 ff.

[168] Suggestio Legatorum ad Hormisdam. Thiel, *Epistolae Romanorum pontificum*, I, p. 874.

[169] M. G. H., *Epist.*, I, p. 263 ff.

[170] Jordan, *op. cit.*, p. 624.

[171] De Rossi, *Inscriptiones*, II, p. 134.

As mentioned before, the official name of the church from at least 431 through the VIII century, was "ecclesia Apostolorum" that is, it was officially consecrated to Saint Peter and Saint Paul. Yet as early as 501, a new evidently unofficial name appears in addition to the official name. In this year for the first time the church is mentioned under the name "a vincula sancti Petri apostoli",[173] in an account of the slaying of one of its presbyters in a street riot. The same name appears in 532/535 in an ex-voto inscription which is preserved in the church and which was dedicated "BEATO PETRO AP. PATRONO SUO" by "A VINCULIS EIUS SEVERUS PB".[174] Again the name is used, when in 544 Arator's poem in honor of the two Apostles is publicly read in "Ecclesia beati petri quae vocatur Ad Vincula . . .".[175] In the middle of the VII century, the Itinerarium Salisburgense calls the church ". . . vincula Petri",[176] and, in the late VIII century, the Itinerarium Einsiedlense mentions it as "Sci. Petri ad Vincula".[177]

Yet it appears that this is merely a popular name. It occurs in itineraries, in ex-votos, in sepulchral inscriptions,[178] in reports on current events, but never in official documents. It seems to have been the current popular name by which the church was known to everybody in Rome.[179] Yet its official name was and remained "titulus Apostolorum". Thus it appears in the signatures of the Roman synods of 499 and of 595.[180] Then, in 600, a new name appears together with the old one; the same two presbyters, Andromacus and Agapitus, who in 595 had signed for the "titulus Sanctorum Apostolorum", are mentioned in a letter of Gregory the Great, dated October 5, 600, as presbyters of the "titulus Eudoxiae".[181] Evidently the name, Eudoxia, in the inscription in the church had led to the belief that she had actually founded the church, thus contributing to the creation of a new name.

---

[172] See the following note.

[173] Duchesne, *L. P.* I, p. 261.

[174] *Ibidem*, I, p. 285, n. 1.

[175] Migne, *Patrologia latina*, Vol. LXVIII, col. 55.

[176] G. B. De Rossi, *Roma Sotterranea*, I, Rome, 1864, I, p. 149.

[177] R. Lanciani, "L'Itinerario dell'Anonimo Einsiedlense", in: *Monumenti Antichi . . . dell'Academia dei Lincei* . . ., I (1892), p. 1 ff.

[178] For instance, the lost sepulchral inscription quoted by F. Cabrol and H. Leclercq, *Dictionnaire d'archéologie et de liturgie chrétienne*, III, 1, Paris, 1903 ff., col. 4, after P. Monsacrati, *De sacris catenis*, Rome, 1750, which Monsacrati had read at S. Lorenzo f.l.m.: "LOCUS SEBASTIANI PB. A VIN".

This new name is used from the VII through the VIII centuries. Whether it was ever an official name, in the proper sense of the word, can hardly be decided. At any rate, it is never used separately after the date of the letter of Gregory the Great. It is used by Hadrian I in combination with both the old official and the popular name "titulus Apostolorum quae appellatur Eudoxiae ad vincula",[182] or in combination with the popular name only, "in titulo Eudoxiae videlicet beati Petri apostoli ad vincula".[183] Only in the biography of Leo III it occasionally occurs alone,[184] but usually in combination with the popular name.[185] Evidently the hitherto popular name received a more and more official character. From 816 on, it appeared by itself in more or less official documents, for example, in the biography of Stephen V,[186] and only once Gregory IV returns to the older combination "ecclesia Apostolorum ad Vincula".[187] But generally the name of the two Apostles is more and more replaced by that of Saint Peter alone, first unofficially and later officially, until the present name, S. Pietro in Vincoli, remains as the only one. And it may well be that the attribution of the chains reflects the same gradual replacement of both the princes of the Apostles by Saint Peter.

---

[179] Change of names of that kind or the contemporary use of both official and popular name have always been and are still frequent in Rome. For instances, see Ch. Hülsen, *Le Chiese di Roma*, Firenze, 1927, passim.

[180] See above, n. 130 and 131.

[181] M. G. H., Epist., II, p. 275.

[182] Duchesne, *L. P.*, I, p. 508.

[183] *Ibidem*, I, p. 512.

[184] *Ibidem*, II, p. 3, 11.

[185] *Ibidem*, II, p. 19.

[186] *Ibidem*, II, p. 49, 50.

[187] *Ibidem*, II, p. 76.

Summing up: It seems to be sound to assume that the church, which had been started by Philippus and which was dedicated by Sixtus III, was consecrated to the princes of the Apostles. Yet from as early as 501, an unofficial replacing of the two Apostles by Saint Peter alone occurred until the name of the latter was supreme. It seems worthwhile to ask whether there was any particular reason for this change.

It is remarkable how Sixtus III, in his dedicatory inscription, seems to apologize to one of the two princes of the Apostles for having him share the honor of the dedication with his fellow apostle.[188] There was no reason for such an apology unless one of the two apostles had a prior right to the church. Such an older right could be connected only with the title of the older ecclesiastical building which preceded the church of Philippus. As a matter of fact, the Hieronymian Martyrology[189] quotes under August 1, the day of the festival of the chains of S. Pietro in Vincoli, a "dedicatio (primae) ecclesiae a beato Petro constructae". Since the main body of the Martyrology dates prior to the papacy of Sixtus III, to whom the latest revisions are due,[190] this can mean only that there existed an ecclesiastical building in Rome which, traditionally, was connected at a very early time with the name of Saint Peter and the consecration of which fell on the same day on which, after the time of Sixtus III, the annual main festival of S. Pietro in Vincoli occurred.[191] The logical conclusion would seem to be that this ecclesiastical building connected with the memory of Saint Peter was the one that preceded the church of Sixtus III. This would explain the apologetic tenor of the dedicatory inscription in the new church, where Saint Peter now had to share honors with Saint Paul, although originally he had been the only patron of the preceding building. It would likewise explain the early appearance of the popular name S. Pietro in Vincoli. Public opinion evidently did not become accustomed to the new dedication to the two Apostles and clung to the old one, to Saint Peter alone, until finally Saint Peter remained victorious.

e. This hypothesis seems to be supported by an inscription which, evidently during the Middle Ages, was to be read in the church. It is quoted only in the XVII century by Martinelli, not from the original which at this time had evidently disappeared, but from some older source:[192]

---

[188] See above, p. 397.

[189] L. Duchesne and G. B. De Rossi, Martyrologium Hieronymianum, Brussels, 1894, in: *Acta Sanctorum, Novembris*, tomus II, p. (99).

[190] L. Duchesne, *Christian Worship*, London, 1904, p. 290.

[191] Kirsch, *op. cit.*, p. 47.

Hoc Domini templum Petro fuit ante dicatum,
Tertius Antistes Sixtus sacraverat olim
Civili bello destructum post fuit ipsum,
Eudoxia quidem totum renovavit ibidem
Pelagius rursus sacravit Papa Beatus,
Corpora sanctorum condens ibi Machabaeorum,
Apposuit Petri pretiosa Ligamina ferri,
Illustris mulier quae detulit Hierusalem.
Et quibus est Petrus, Neronis tempore vinctus,
Augusti mensis celebrantur festa calendis.
Huc accedenti purgantur crimina cuncta.

The first two lines of this inscription seem to state that the church had originally been dedicated to Saint Peter and was later consecrated by Sixtus III. It does not say that it was consecrated to the Apostles, but the fact that a second consecration was undertaken by Sixtus is very definitely stressed.

Unfortunately, however, the date of this inscription is very uncertain. It certainly is at least as late as the VI century, as is evident from the reference to a pope Pelagius and to the relics of the Maccabeans.[193] But it does not occur in any of the VII, VIII or IX century syllogæ.[194] The syllogæ, of course, are far from complete and the inscription might very well be late VI century. If later, it might have repeated a reliable early tradition. Thus, although at the present state of our knowledge, the inscription cannot be relied upon, in view of the conclusions drawn from the other sources discussed above, it may prove a supplementary corroboration of the hypothesis that the edifice preceding the church of Philippus and Sixtus was dedicated to Saint Peter.

To summarize briefly the results of the documentary evidence, it becomes obvious that the church was started before 431 by Philippus and that it replaced an older building of unknown form. This older building was dedicated in all likelihood to Saint Peter. The new church, consecrated by Sixtus III, possibly in the presence of Eudoxia, in 439–440, and certainly supported by Imperial donations, was dedicated to the Apostles Saint Peter and Paul. This documentary evidence seems to tally with the evidence given by the structural analysis of the building. This analysis had shown that the transept and the apse were earlier than the nave, although forming part of one and the same building campaign. The existence of different successive projects in the transept suggests even a building activity of quite some duration in the eastern parts of the edifice. On the other hand, the lower parts of the West façade must have been erected by 439–440 when they were decorated with the dedicatory inscription of Sixtus III. Only afterward the clerestory of the nave seems to have been erected, and the changing distances between the windows in this clerestory as well as their changing width go to prove that also the process of completing the nave took quite a long time. The proportions of this clerestory suggest a date about or shortly after the middle of the V century. Thus one will be safe in assuming that the building was started about 420 and terminated about 450.

---

[192] Martinelli, *op. cit.*, p. 248 f.

[193] Cabrol-Leclercq, *op. cit.*, III, 1, col. 3 ff., assumed that the relics were transferred first from Antioch to Constantinople and then, under Pelagius I (556–561), to Rome. Since, however, the relics were still at Antioch when Antoninus of Piacenza visited that city after 570 (as pointed out by Cabrol-Leclercq, *op. cit.*, I, 2, col. 2378 f.) the translation seems to have taken place under Pelagius II (579–590).

[194] Marucchi, *op. cit.*, p. 317, must be mistaken when assuming that the inscription appears in the Mirabilia.

V

As a whole, the church of S. Pietro in Vincoli seems to present itself as a normal Early Christian basilica of the well-known Roman type. Yet there is one element in the church which is different from the normal and which is of the utmost importance. It is the transept, with its clear division into a center bay and wings. This division, as one recalls, was at first intended to be achieved by walls, each with three openings in their upper parts and with arcades, possibly on columns, below. Later, although during the same construction period, the arrangement was simplified by eliminating these upper walls and by retaining only four pilasters in their places. Whether or not the lower parts of these pilasters were connected by arcades cannot be ascertained (Fig. 14). Yet in the first as well as in the second project the transept of S. Pietro in Vincoli belongs to a transept type which plays a considerable role in Early Christian architecture. We would like to discuss this particular transept type in greater detail since to my knowledge this has never been done before.

Leaving aside the question of the origin of the transept as such,[195] it is nevertheless certain that transepts appear in the Early Christian basilicas of Rome from the early IV century on and that from the early V century they appear also in the Near East. But certainly, as Kirsch and Schneider[196] have pointed out, different types of transepts should be distinguished from one another. Evidently these transepts are in type independent of each other, not only from a structural, but also from a liturgical point of view, and very different in origin. There is first the continuous narrow transept, a structure which equals the nave in height, which is undivided, and which is interposed between apse and nave. In this form the transept is a semi-independent structure, and as a rule it seems to be erected over the tomb of a Saint or possibly over a famous relic. It appears to be a memorial building, structurally connected with, yet in purpose and in religious significance opposed to, the congregational hall, the nave with its aisles. While this latter area was reserved for the congregation, the transept was reserved for the clergy and for the cult of the martyr.[197] One might even ask whether the transept in this form is anything but a substitute for the circular or octagonal memoria which formed an indispensable part of the Constantinian churches of the Near East, either as an independent structure, as in the Holy Sepulchre at Jerusalem,[198] or directly attached to the nave, as in the Church of the Nativity at Bethlehem.[199] This continuous semi-independent transept with memorial character seems to occur with particular frequency in Rome.[200] Chronologically, it seems to be limited, in Early Christian architecture proper, to the IV century. It appears at St. Peter's about 325, at St. John in the Lateran about 335, at St. Paul's, which was certainly copied from St. Peter's, after 384, and possibly in the first church of S. Lorenzo f.l.m. about 390. Later it occurs only once, between 625 and 638, at S. Pancrazio, until, in the early IX century, it becomes a common feature of Roman churches and of Carolingian churches in general. It is evident that, with the exception of St. John in the Lateran, the churches to which these transepts belong are, without exception, basilicæ ad corpus, erected in connection with the tomb of a saint. While this type seems to be generally limited to the west, it occasionally occurs also outside Rome, for example, at S. Maria di Capua Vetere;[201] it may have found its way even to Asia Minor in some isolated examples such as basilicas E and E′ at Sagalassos, judging as far as the sketchy groundplans published allow.[202]

---

[195] J. P. Kirsch, "Il Transetto nella basilica cristiana antica", in: *Scritti in onore di Bartolommeo Nogara*, Rome, 1937, p. 205 ff. (quoted as Kirsch, "*Transetto*"), gives an extraordinarily clear and complete survey of the older and more recent theories on the origin of the transept. His survey makes it unnecessary to repeat it here.

[196] A. M. Schneider, *Die Brotvermehrungskirche von et-tâbga*, Paderborn, 1934, p. 35 ff.

[197] Th. Klauser, Die konstantinischen Altäre der Lateransbasilika, in: *Römische Quartalschrift*, 43 (1935), p. 179 ff., attempts to derive the transept from the necessity of creating room for the seven altars in the basilica of the Lateran. This theory does not explain the spreading of the transept to buildings where such a number of altars did not exist. Schneider, *op. cit.*, p. 36, has pointed out very clearly the different elements involved in the origin of the transept, the structural elements as well as its connection with the martyr cult. He does not emphasize, however, the parallel with the memorial buildings of the East nor does he point to the difference in origin and purpose of the different transept types.

[198] H. Vincent and F. M. Abel, *Jérusalem Nouvelle*, II, 1, Paris, 1914, passim.

This type is different from a second type of transept which is as wide, or almost as wide, as the nave. Like the nave, it is accompanied by aisles which frequently continue along the short sides so as to surround the wings on three sides. In the center, where nave and transept intersect, the tomb of the Saint is placed. In other words, the transept is not a semi-independent structure, attached and at the same time opposed to the nave, but it is united and architecturally equivalent to the nave. It is not a memorial building which shelters the tomb of the Saint and is merely attached to the congregational hall; together with the nave, it radiates from the center formed by the tomb of the Saint towards three sides, while the fourth side is usually occupied by the apse. The tomb forms the religious as well as the architectural center of the whole edifice and consequently a radiating element dominates buildings of this type even in its latest descendents.[203] The so-called Arcadius basilica at the Menas city [204] is the best example of the type (Fig. 15).

---

[199] W. Harvey, "The early basilica at Bethlehem," in: *Palestine Exploration Fund, Quarterly Statement*, 1936, p. 28 ff.; H. Vincent, "Bethléem", in: *Revue Biblique*, 45 (1936), p. 544 ff.; 46 (1937), p. 93 ff.

[200] Kirsch, *Transetto*, has attempted to limit its occurrence to basilicas with a nave and four aisles.

[201] G. Chierici, in: *Atti del III Congresso di Archeologia Cristiana*, Rome, 1934, p. 204 ff., Fig. 1.

[202] K. Graf Lanckoroński Brzezie, G. Niemann, E. Petersen, *Städte Pisidiens und Pamphyliens*, II, Vienna, 1890–92, p. 149 f., Fig. 123 and p. 152 f., Fig. 126; H. Rott, *Kleinasiatische Denkmäler*, Leipsic, 1906, p. 14 ff., Fig. 17. The churches of Epidauros and Paramythia, which Schneider, *l.c.*, enumerates in this group, have no connection with it, nor does the single-naved church of Kyzyllyk, which was published by R. Paribeni and P. Romanelli, "Studii e ricerche archeologiche nell'Anatolia meridionale", in: *Monumenti Antichi dell'Accademia dei Lincei*, XXIII (1913), col. 267 ff.

The derivation of this plan of transept can undoubtedly be traced back to a well-known Early Christian type of the IV century, the cross-shaped memoria. In these edifices, four wings of equal length radiate from a center room which contains the tomb of the Saint or the venerated object.[205] A building of this simplest type, dated 379–80, is preserved at Antioch-Kaoussié,[206] while another one is known to have existed at the Well of Jacob near Sichem.[207] By the early V century this type is further developed.[208] The wings are transformed to consist of nave and aisles and, instead of all being equal, one of the wings is distinguished by the addition of an apse.[209] Thus the longitudinal axis is slightly more emphasized than the transverse of the transept wings. This type existed probably in the church founded by Eudoxia the Elder at Gaza in 401; [210] it has been excavated in the Church of the Martyrs and Apostles at Gerasa (464–65) [211] and in the cross church at Salona (490–

---

[203] Schneider, *op. cit.*, p. 37, has pointed to the differences of this type, which he calls type II, from the one discussed before. I disagree, however, with his interpretation of the second type. According to him, the nave in this type would be through-running and the transept would be attached to it. He overlooks the centralizing and radiating element; he is bound to do so because he does not take into consideration the difference in origin and purpose.

[204] C. M. Kaufmann, *Die Menasstadt*, Leipsic, 1910; F. W. Deichmann, "Zu den Bauten der Menasstadt", in: *Jahrbuch des Deutschen archäologischen Instituts, Archäologischer Anzeiger*, 52 (1937), col. 75 ff.

[205] For this problem see S. Guyer, "Der Dom von Pisa . . ." in: *Münchner Jahrbuch der bildenden Kunst*, 10 (1933), p. 351 ff., particularly p. 360 ff.

[206] *Antioch-on-the-Orontes*, II, *The excavations, 1933–36*, *Publications of the Committee for the Excavation of Antioch*, ed. R. Stillwell, Princeton, 1938, p. 5 ff.

[207] Its plan is preserved in a VII century drawing by Arculf; see Arculfus, *De locis sanctis*, lib. II, cap. 21, ed. Migne, *Patrologia latina*, Vol. LXXXVIII, col. 802.

[208] Distantly related, although on a much smaller scale, are the cross churches with barrel vaulted wings in Anatolia as published by Rott, *op. cit.*, passim, for example, p. 163, 182 ff., 188 ff., and the cross churches with stone slab covered wings in the Hauran, discussed by J. Lassus, in: *Atti del III Congresso*, p. 479 ff.

[209] Guyer, *op. cit.*, p. 360 ff., and idem, in: *Atti del III Congresso*, p. 433 ff., as well as J. W. Crowfoot, *Gerasa*, New Haven, 1938, p. 190 ff., have carefully studied the material pertaining to this group.

[210] This church is known only through descriptions; Crowfoot, *op. cit.*, p. 192 ff., and R. W. Hamilton, "The churches of Gaza", in: *Palestine Exploration Fund, Quarterly Statement*, 1930, p. 190 ff.

[211] Crowfoot, *op. cit.*, p. 256 ff.

540).[212] Kalat Siman, of about 450, belongs to the same group, but it has an octagonal rather than a square center bay which was probably covered by a wooden dome.[213] The pre-Justinian church of Saint John at Ephesos, unfortunately not dated, transforms this almost pure cross type still further by more strongly emphasizing the W.–E. axis through the addition of a fore-church in the west and by increasing the number of aisles in the E. arms to four.[214] Thus the transverse cross arms become a kind of transept. Yet Ephesos still retains the characteristic elements of a cross memorial with the tomb of the Evangelist forming the structural and cult center of the whole building. The V century church of H. Demetrios at Salonica, which preceded the present one, shows the same stage of the development. This time, apparently, a rectangular choir arm was surrounded by aisles, like the nave and the cross wings, yet, like the cross wings, considerably shorter than the nave.[215] The substitution of a regular semicircular apse for such a chancel arm resulted

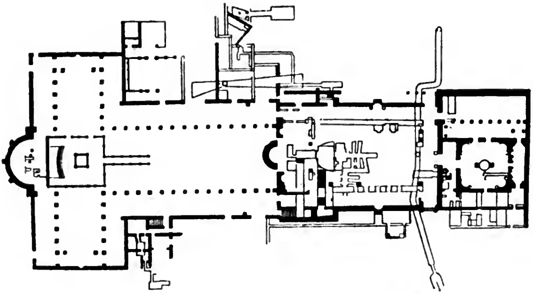


FIG. 15. Menas city, Arcadius basilica, plan (O. Wulff, *Altchristliche und Byzantinische Kunst*, 1914, I, Fig. 226).

---

[212] R. Egger and W. Gerber, *Forschungen in Salona*, I, Vienna, 1917, p. 23 ff., 91; W. Gerber, *Altchristliche Kultbauten in Istrien und Dalmatien*, Vienna, 1912, p. 120.

[213] H. C. Butler, *Early Churches in Syria*, Princeton, 1929, p. 97 ff. (quoted as Butler, "*Early Churches*"); O. Krenker, *Die Wallfahrtskirche des Simeon Stylites in Kal'at Sem'ân, Abhandlungen der Preussichen Akademie der Wissenschaften*, 1938, *Phil. Hist. Klasse*, 41, Berlin, 1939.

[214] J. Keil, "Vorläufiger Bericht über die Ausgrabungen in Ephesos", in: *Jahreshefte d. Oesterreichischen Archäologischen Instituts*, Beiblatt XXVII (1932), col. 61 ff.

[215] The structural analysis of H. Demetrios at Salonica is exceedingly complicated. I am basing my statement that there existed an earlier building of the V

in creating the type of the Menas basilica, the date of which is not quite certain [216] (Fig. 15). The derivation of its strange transept becomes now manifest: the cross-shaped memoria of the IV century, enlarged by aisles, has been changed by transforming the E. arm into an apse, the W. arm into a nave, and the cross wings into what might be called a "cross transept". Evidently this transformation took place in the course of the V century in the Coastlands of the Near East where the cross-shaped memoria had originated. With the exception of Ephesos, which is undated, the edifices of the group are all of the V century. Evidently the transformation was intended as an adaption of the cross-shaped memoria to the basilica with a continuous transept as represented by the IV century basilicas of Rome. Still, the prototype, the cross memoria, remains obvious even in the latest descendents of the type. The radiating centralizing element is never lost; the transept wings are always accompanied by aisles, although sometimes in a rudimentary form.[217]

These rudimentary derivatives spread during the V and possibly the VI centuries through the Near Eastern Coastlands, including Asia Minor, Palestine, and Bulgaria.[218] Among them are the two basilicas at Perge [219] and the basilica of Tropaion,[220]

---

century, which would be connected with the date of a foundation in 412–13, and its reconstruction largely on my own observations. See also G. Soteriu, "He Basilike tou Agiou Demetriou . . .", in: *Deltion Archaeologikon Symplegma*, IV (1918), p. 1 ff.; idem, "Hai palaiochristianikai basilikai tes Hellados", in: *Ephemeris Archaeologike*, 1929 (1931), p. 211 f. (quoted as Soteriu, "*Ephemeris, 1929*"); Schneider, *op. cit.*, p. 35 ff.

[216] Kaufmann, *op. cit.*, dates it 400, Deichmann, *op. cit.*, at some later time in the V century.

[217] This analysis is definitely opposed to Schneider's analysis, *op. cit.*, p. 35 ff., who interprets the type as a continuation of the nave into and through the transept, an interpretation which disregards the structural characteristics of the building as well as the dominating position occupied in it by the tomb of the Saint or the venerated object. Based on this erroneous analysis and a somewhat indiscriminate selection of material, Schneider is also bound to confuse the derivatives of the type and to enumerate side by side as members of the same group churches, such as the Arcadius basilica of the Menas city and H. Demetrios at Salonica, which actually belong in this group; Nikopolis A and B and the "transept basilica" extra muros at Korykos, which are in an entirely different group which we will discuss below; Dodona, which is again slightly different; the Tschanli Kilisse at Gereme, a small cross-shaped church of the Anatolian type discussed before; and finally Henchir-el-Guesseria in North Africa and Serdjilla in Syria, both of which have not a transept but enlarged pastophories.

[218] Schneider, *op. cit.*, p. 38 f., speaks of these derivatives as of "mixed types".

[219] Rott, *op. cit.*, p. 46 ff., Figs. 19 and 21.

[220] W. Liesenberg, *Der Einfluss der Liturgie auf die frühchristliche Basilika.* Neustadt, H., 1928, p. 143: R. Netzhammer, *Die christlichen Altertümer der Dobrogea*, Bukarest, 1906, p. 192 ff.

where the aisles accompany the transept wings only along the W. side and along the N. and S. sides respectively,[221] while they are missing on the E. side. At et-Tabhgha [222] this type has been simplified even more by omitting the aisles even on the N. and S. sides of the transept, thus preserving them only along the W. side, but adding a screen of columns in the opening between nave and transept.

The transept of S. Pietro in Vincoli is entirely different from both the Roman continuous transept and the cross transept. It does not form a continuous semi-independent structure which opposes itself to the nave as do those in the Roman basilicas of the IV century and in their immediate descendents. Nor does it consist of two wings surrounded or accompanied by either complete or rudimentary aisles, which would make the transept resemble the nave and create the impression of three similar wings radiating from a center, as in the cross transept churches of the Near Eastern Coastlands. At S. Pietro in Vincoli, the transept consists of three clearly differentiated parts, the center bay and the two transept wings. The center bay forms a transverse oblong, the wings, two longitudinal ones. The center bay seems to be a continuation of the apse towards the nave, a kind of forechoir, an elongated chancel; the transept wings seem to flank this chancel right and left. The center bay is clearly marked off against the nave as well as against the transept wings. This analysis holds for both the first unabbreviated stage, with dividing upper walls, and for the second one, without them; it also holds if the reconstruction of the last stage is made without any arcades between the four projecting piers. It is a form of transept all by itself and quite independent of either the Roman continuous transept of the IV century or the Near Eastern cross transept of the V century. One might call it a "tripartite transept" (Fig. 14).

S. Pietro in Vincoli, however, is not an isolated representative of this type of transept. There exists a large group of churches with transepts of the type of S. Pietro in Vincoli or, to put it more precisely, of the different types which have left their traces at that church: one with a two storied partition wall between center bay and transept wings; one with only an arcade between them which may have succeeded this first one; and one in which they were possibly separated merely by projecting pilasters.

---

[221] In basilica A at Perge and at Tropaion, this arrangement is well preserved; in basilica B at Perge, it can be reliably reconstructed.

[222] Schneider's date of 375 can hardly be maintained any longer in view of his own discovery of a smaller chapel below the present church; see: "Neue Funde in et tabga", in: *Oriens Christianus*, XXXIV (1937), p. 59 ff. Also the character of the mosaics of the large church compares rather to the V century mosaics at Gerasa than to any earlier ones.

This latter type is represented by a group of basilicas in Greece, the basilica A of Nikopolis,[223] the basilica at the Ilissos near Athens,[224] and another one at Arkitsa-Daphnousia near Lokris.[225] In all of these churches the nave is accompanied by two aisles, and in all of them the transept is separated, by means of protruding piers, into a center bay and two wings. In basilica A of Nikopolis the cross form of the piers between nave and transept seems well established[226] (Fig. 16). In the Ilissos ba-

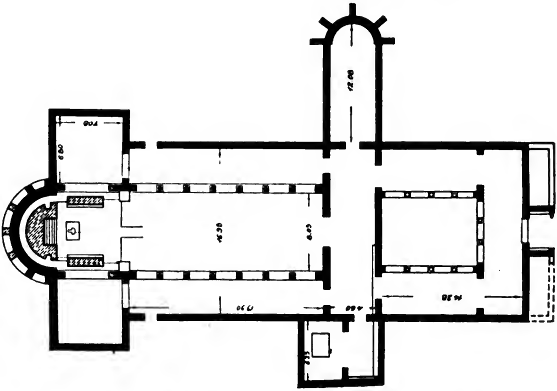


Fig. 16. Nikopolis A, plan (Soteriu, *Ephemeris*, 1929, Fig. 37).

---

[223] A. Philadelpheus, "Anaskaphai Nikopoleos", in: *Ephemeris*, 1916, p. 33 ff., 65 ff.; 1917, p. 48 ff.; 1918, p. 34 ff.; Soteriu, *Ephemeris*, 1929, p. 206 f.; A. Orlandos and G. Soteriu, "Anaskaphai Nikopoleos", in: *Praktika Archeologikes Hetairias*, 84 (1929), p. 83 ff.; 92 (1937), p. 78 ff.; 93 (1938), p. 112 ff.

[224] G. Soteriu, "Palaiachristianike Basilike Ilisou", in: *Ephemeris*, 1919, p. 1 ff.; 1929, p. 208 ff.

[225] A. Orlandos, "Une basilique paléochrétienne en Locride", in: *Byzantion*, V (1929/30), p. 207 ff.; Soteriu, *Ephemeris*, 1929, p. 207 f.

[226] Soteriu, *Ephemeris*, 1929, p. 206, Fig. 37; *Praktika*, 92 (1937), p. 78 ff.; 93 (1938), p. 112 ff. In earlier publications the piers are indicated as being T-shaped with free standing columns, carrying the triumphal arch, a reconstruction disproved by later excavations.

silica only strong, square foundation walls are preserved instead of these piers, yet the shape of these very foundation walls makes it clear that the piers must have been cross shaped.[227] In both these churches, comparatively small openings lead from the aisles into the transept wings. At Nikopolis it is a single arch, almost like a door; in the Ilissos church it is a double arch, yet with rather narrow openings. In both these churches the center bay of the transept was segregated from the wings and from the nave by choir screens, remnants of which have been excavated. At Nikopolis a set of beautiful mosaics covers the wings of the transept and the nave, while the center bay of the transept and the aisles evidently had a plain pavement. The few remnants in the Ilissos basilica point to a similar arrangement of the pavement decoration. In the Ilissos church the center bay forms a wide oblong and the wings long rectangles, thus bringing the short walls of the transept into line with the walls of the aisles. Along the short sides of the transept and the E. end of the aisles, other narrow rooms are added, possibly lower parekklesiai. At Arkitsa this arrangement has been slightly changed: the center bay is wider than the nave and the wings protrude slightly beyond the aisles, thus creating a plan which corresponds perfectly to that of S. Pietro in Vincoli. Finally, at Nikopolis A the center bay, as well as the transept wings, is almost square and the wings protrude considerably beyond the walls of the aisles (Fig. 16).

It is evident that these three churches form one closely related group, and that, in turn, they are closely tied up with S. Pietro in Vincoli in its final stage. The existence of a tripartite transept is evident here and there, and it is likewise evident that in S. Pietro in Vincoli, as well as in this group, the oblong center bay is marked off as a kind of forechoir arranged in front of the apse. Of course, it cannot be established whether, in the Greek churches, the piers, which protrude from the corners of the center bay, terminated straight at the top of the transept wall, as was the case at S. Pietro in Vincoli, or whether they carried an arch which would have been either as high or slightly lower than the triumphal arch. Whether one or the other solution is chosen for the reconstruction, the type at any rate closely resembles S. Pietro in Vincoli in its final stage, unless at S. Pietro in Vincoli there were still in this final stage arcades connecting the E. and W. piers and thus accentuating the separation of the center bay from the wings.

---

[227] There seems to be no reason for assuming the existence of a dome over the center bay as is done by Soteriu, "*op. cit.*", *Ephemeris*, 1919, p. 1 ff.

Churches with arcades of this type, either without or with upper walls with openings and thus resembling S. Pietro in Vincoli in its first project, existed also in Greece. An edifice of this plan has been excavated, for example, in basilica B at Nikopolis (Fig. 17).[228] The nave, instead of being flanked by only two aisles, has four. But, as at S. Pietro in Vincoli and at Nikopolis A, it runs into a tripartite transept which, however, does not protrude beyond the outer aisles. Of course, it cannot be ascertained whether the transept was really as high as the nave and whether its wings were as high as its center bay. Yet it is certain that there existed a center bay and two transept wings; that the center bay formed a transverse oblong while the wings were longer than wide; and that the wings were separated from the center bay by an arcade of three arches which were carried by two square marble piers. There is no way of determining exactly the shape of the piers at the corners between nave and transept. The excavation did not lead to any definite results on that point. What is well preserved is the arrangement of the choir screens, and this arrangement reveals clearly the use of the different parts of the transept. The center part

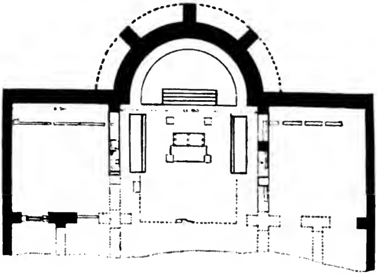


FIG. 17. Nikopolis B, plan (Soteriu, *Ephemeris*, 1929, Fig. 33).

---

[228] G. Soteriu and A. Orlandos, "Anaskaphai Nikopoleos" in: *Praktika Archaeologikes Hetairias*, 85 (1930), p. 79 ff.; A. Philadelpheus, in: *Deltion Christianikes Hetairias*, IV (1928), p. 46 ff.; Soteriu, *Ephemeris*, 1929, p. 200 ff.

is completely taken up by the choir proper, while in the wings choir screens run at a distance of only 1.90 m along their E. walls. The whole choir screen thus results in a strongly projecting center part and in receding lateral parts in the wings, an arrangement very similar indeed to the one traced at S. Pietro in Vincoli.[229] Other choir screens, one of them with an entrance clearly preserved, separate the transept wings from the aisles.

Closely related to this basilica B at Nikopolis is a church excavated at Epidauros.[230] Again the nave is flanked by two aisles on either side, but the transept is divided into five instead of three parts, an oblong bay and, as it were, four transept wings, two on either side. The center bay and the inner wings are separated by a double arcade supported by one column in the middle; the two outer wings are separated by triple arcades resting again on columns. Also significant is the arrangement of the choir screens and, closely connected with it, the distribution of the mosaic floors. The apse, the center bay, and the inner wings of the transept evidently belong together since they are separated from the nave and from the inner aisles by screens and are laid out with large flagstones. The outer wings connect with the outer aisles only by doors and have a pavement of small flagstones. Only the floor of the nave is covered with a mosaic.

A very similar arrangement existed a third time, in the basilica which the American School at Athens has uncovered at Corinth.[231] Again the transept is clearly separated from the nave, again the center part is wider than it is long, while the wings are longer than they are wide. Again center bay and wings are separated from one another by arcades, although at Corinth these arcades consisted of only two arches resting on one center pier. Again the choir screens mark off the whole center bay as an elongated chancel. The piers between nave and transept are definitely cross shaped, exactly like those at S. Pietro in Vincoli and contrary to those in the church in Epidauros where they seem to have been oblong in shape. Unlike

---

[229] See above, p. 392 f.

[230] P. Kabbadia, "Anaskaphai en Epidauro", in: *Ephemeris*, 1918, p. 172 ff.; Soteriu, *Ephemeris*, 1929, p. 198 ff.; idem, He basilike tes Epidaurou [illegible] *Praktika Akademias Athenon*, IV (1929), p. 91 ff.

[231] R. Carpenter, "Researches in the topography of ancient Corinth", in: *American Journal of Archæology*, II ser., 33 (1929), p. 345 ff.; Soteriu, *Ephemeris*, 1929, p. 127 f.; J. De Waele, in: *Atti del III Congresso*, p. 371 ff.

Epidauros and Nikopolis B, the nave is flanked by only two aisles, but, as in these churches, the aisles open in a double arcade towards each of the transept wings. Two small chambers east of the transept wings, evidently pastophories, completed the plan. The manner in which they are joined to the transept walls makes it slightly doubtful whether they belong to the original building.[232]

This second group of churches in Greece separates the center bay from the transept wings by means of arcades, as must have been the case in the first project of S. Pietro in Vincoli and was possibly still a feature retained in a second stage which had discarded the upper walls with their openings. Whether any of these churches in Greece had upper walls, with openings, rising from the arcades, as was the case in the first project of S. Pietro in Vincoli, cannot, of course, be ascertained since only their foundation walls were found; nor can it be established whether such upper walls, if they existed, were pierced by openings as was the intention in the first project of S. Pietro in Vincoli.

It is quite evident that these two groups of churches are closely related to one another and that S. Pietro in Vincoli, in turn, is closely related to both these groups in its successive stages. Yet certain differences between the two Greek groups, although only slight, might be worth mentioning. While the group with arcades between center bay and transept wings, as a rule, has four aisles, the other group has only two. While in the first group, the transept wings usually do not project beyond the aisles, they do so in the second group, except in the Ilissos basilica. These differences may become important in discussing the origin of the tripartite transept.

Unfortunately, neither the first nor the second group of these basilicas in Greece is definitely dated, although it is generally agreed [233] that most of them are of the V century. At Nikopolis A, the style of the mosaics, as well as their inscriptions which refer to a bishop Dumetios, seem to point to the second half of the V century.[234] In basilica B at Nikopolis an inscription in a small lateral chapel mentions a bishop Alkision and this gives the years 491–518 as a terminus ad for the chapel [235] and consequently as a terminus ante for the main church. Arkitsa is dated by Orlandos [236] in the early V century; Corinth by de Waele [237] in the late IV or early V century.[238] The Ilissos basilica may be either from the beginning of the V century or possibly a few decades later.[239] At any rate, the V century seems to include all of the buildings in this group with the possible exception of the church in Dodona [240] which may belong to the middle of the VI century. The date of the Greek churches would, therefore, be in complete conformity with that of S. Pietro in Vincoli. They cover a territory reaching from the Peloponnesos to Epirus and evidently form one of the common types of Early Christian architecture in Greece during the V century.

---

[232] The basilica excavated at Dodona, Soteriu, *Ephemeris*, 1929, p. 203 f., belongs in the same group with the one difference that the transept wings end in apses thus giving the E. end of the edifice the shape of a triconch. Soteriu is probably right in dating the edifice into the period of Justinian, about the middle of the VI century. It seems to be a late derivative of the type discussed.

[233] Soteriu, *Ephemeris*, 1929, passim.

[234] Philadelpheus, "*op. cit.*", in: *Ephemeris* (1916), p. 65 ff., and Soteriu, *Ephemeris*, 1929, p. 206 ff.

Outside Greece only one isolated instance of such a church is preserved in church I at Bosra, in Syria.[241] The forechoir forms a transverse oblong, the transept wings long receptacles. The main piers are cross shaped. It remains doubtful whether the wings were separated from the forechoir by a double arcade [242] or merely opened in one wide and tall arch towards it. As in Corinth, small pastophories are added to the E. walls of these transept wings. Yet a uniform group of the type seems to exist only in the Greek churches and at S. Pietro in Vincoli. The latter is undoubtedly closely tied up with this Greek group of buildings, and since its transept type is unique in Rome, while common in Greece, there is hardly any reason to doubt that the type was imported into Rome from the Greek peninsula. Since Greece was under the jurisdiction of the Roman See from the early IV century to the end of the V,[243] it is not astonishing to find such close interrelations.

---

[235] *Ibidem*, p. 201, n. 1.

[236] Orlandos, *op. cit.*, p. 207 ff.

[237] De Waele, *op. cit.*, p. 380 f.

[238] The remnants of a capital found in the little memorial chapel to the right of the right aisle are dated by R. Kautzsch, *Kapitellstudien*, Berlin and Leipsic, 1936, p. 130 f., not into the early V century but as late as the last quarter of the V century. Since we do not know anything about the original place of the capital this need not change the date of the church. The other capitals which might have come from this basilica, now at the Museum of Corinth, have actually no definitely known provenance and may come from any of the many churches in Corinth (Kautzsch, *op. cit.*, p. 75).

[239] Soteriu, "*op. cit.*", in: *Ephemeris*, 1919, p. 1 ff., gives it the earlier date.

[240] See above, n. 232.

[241] Butler, *Early Churches*, p. 118 f. The church is not dated. In the latest issue of the *Jahrbuch des Deutschen Archäologischen Instituts*, 54 (1939), *Archäologischer Anzeiger*, mention is made of an excavation near the Top Kapu Saray at Constantinople where remnants of a church were found containing a transept which from the aisles was accessible only by doorways.

[242] As Butler, *l.c.*, assumed; no traces of such a double arcade have been found, however, as Butler himself admits.

Only a suggestion can be made as to the origin of this type.[244] The manner in which the wings are separated from the center bay by piers or arcades, the manner in which they are differentiated from the center bay by the arrangement of the screens and by the layout of the pavements, the way in which they communicate with the aisles, frequently only by doors, and in which sometimes they are separated from them by screens, all manifest that they are intended to form segregated rooms which flank the center bay. On the other hand, the center bay is likewise clearly marked off as a forechoir. This recalls the manner in which in churches in Syria and elsewhere throughout the Near East and North Africa, apse and forechoir are frequently flanked by two small square side chambers, the pastophories.

It would lead much too far to go into the question of the origin of the pastophories or into the question of their development or of their liturgical use; they must have been some kind of sacristies, one of which, the prothesis, was, at least from the V century on, used for preparing the Eucharist. In the IV and far into the V centuries they are frequently just square side chambers which flank the apse and which are accessible only by doors from the aisles.[245] But from the early V century on, new types of pastophories are introduced beside the old ones. In some instances one or both of the pastophories became elongated so as to protrude far beyond the outer walls of the aisles; consequently in groundplan these pastophories sometimes resemble a tripartite transept. Such is frequently the case in North

---

[243] Duchesne, *Christian Worship*, p. 41 f.

[244] This same idea must be in the minds of the Greek archæologists. Whenever they speak of these transept wings, they call them pastophories, but without giving any reasons for it. Evidently to anyone raised in the rites of the Greek church, two rooms flanking a choir are just pastophories.

[245] Syria: Fafirtin, 372; Babiska, 401; Dar Kita, 418; Kirbet Hasan, 507; Kirbet Tezin, 585. Transjordania: Gerasa, Cathedral, about 375; Gerasa, Church of the Martyrs and Apostles, 464–65. North Africa: Timgad, Great basilica, early V century (?); Khamissa, VII century (?). Europe: Salonica, Eski Djuma, early V century; Ravenna, S. Apollinare in Classe, about 530, and numerous others.

Africa[246] and occasionally in Syria.[247] At Serdjilla there is even an anteroom arranged in front of these protruding pastophories. The reason is evidently to create more space for liturgical purposes. In other structures, possibly at a slightly later date, the entrances to the pastophories are altered, again probably in connection with a change in liturgy. They either connect one of the side chambers with the aisle and the other with either the apse alone[248] or with both apse and aisle,[249] or else both side chambers communicate with the aisle as well as with the apse.[250]

At the same time, however, an even more important development takes place; the chancel is elongated so as to create a forechoir in front of the apse. Consequently the pastophories become either elongated, so as to flank not only the apse but also the forechoir, or anterooms are arranged in front of the pastophories so as to flank the forechoir;[251] or the pastophories instead of flanking the apse, are pushed forward so as to occupy the last bay of the aisle and to flank the forechoir only, with which they communicate through doorways.[252] This arrangement with forechoir seems to occur frequently throughout the Christian world, in Syria, in Egypt (Hammam el Farki and East church, Serre),[253] in Greece (Hypate),[254] in the V and VI centuries, and also in Rome, and here at a very early date; at S. Lorenzo in Lucina which is about contemporaneous with S. Pietro in Vincoli, the forechoir is flanked at least on one side by what seems to have been a pastophory.[255]

More important than these different types of pastophories is perhaps a variant of this pattern of an elongated forechoir flanked by side chambers which occurs in a group of churches in Cilicia and elsewhere in Asia Minor. In this group the rooms, which occupy the last bays of the aisles and thus flank the fore-choir, are not secluded by walls as are the pastophories mentioned before. They open towards both the aisles and the fore-choir with wide arches, thus creating an arrangement which in groundplan resembles a tripartite transept with a wide, but rather shallow center bay and square, but small wings. Such is the plan of the cathedral of Korykos which is dated either 429 [256] or 533,[257] in basilica IV at Kanytelideis,[258] in Gül Bagtsche near Izmir,[259] and in the so-called "transept basilica" at Korykos.[260] In Anatolia, basilica XXXII at Bin-Bir-Kilisse [261] shows a similar arrangement and so did the church of our Lady of the Hundred Gates at Paros in its first stage.[262] In both of them the last bay of each aisle is set apart by the use of a pier different from those employed in the nave and is thus marked off as a special unit. Later, in 611, the pattern appeared in Transjordania, in the church of bishop Genesius at Gerasa.[263] It spread towards North Africa where it is found at Henchir el Atech,[264] and possibly in the churches of the White Convent (Der-el-Abiad), about 450, and of the Red Convent (Der-el-Akhmar), a few decades later, both near Sohag in Egypt; in the latter two, the three bays, which are arranged at the end of the nave in front of the triconch, should in all likelihood be interpreted as a forechoir flanked by two wide open "transept pastophories".[265] Again the same arrangement existed in the church erected inside the Erechtheum at Athens in the VI century.[266]

---

[246] Gsell, *op. cit.*, p. 142, gives a list of 19 churches in which either one or both sacristies project beyond the aisles.

[247] Butler, *Early Churches*, p. 130 f. (Hass, South church); p. 67 (Kalota, East church, dated 492); p. 210 (Burdj-Hedar, West church) and others.

[248] Gsell, *op. cit.*, p. 176 (Benian, dated 434–439); p. 207 (Kherbet Guidra, dated 444); Butler, *Early Churches*, p. 91 (Il Karis); p. 122 (Zorah, dated 515).

[249] Frequent in Syria, Butler, *Early Churches*, passim, for example, p. 49 f. (Ksedjbeh, dated 414); p. 47 (Um-is-Surab, dated 489).

[250] Frequent in North Africa, Gsell, *op. cit.*, p. 232 (Morsott); p. 267 (Tebessa) and in Syria, Butler, *Early Churches*, p. 125 (Bosra, dated 512); p. 163 f. (It Tuba, dated 588). The arrangement appears also in Egypt, in the Balkans, and in Asia Minor.

[251] Butler, *Early Churches*, p. 29 f. (Simkhar); p. 59 f. (Serdjibleh).

[252] *Ibidem*, p. 71 ff. (Kalb Louzeh, dated about 480); p. 125 ff. (Bosra, cathedral, dated 512); p. 81, 190 (Il Anderin, St. Michael and St. Gabriel).

[253] E. S. Clarke, *Christian Antiquities in the Nile Valley*, Oxford, 1912, p. 59 ff.

[254] Soteriu, *Ephemeris*, 1929, p. 186 ff.

[255] Krautheimer and Frankl, *op. cit.*, p. 388 ff.

---

[256] E. Herzfeld and S. Guyer, *Meriamlik und Korykos . . .* , Manchester, 1930, passim.

[257] Kautzsch, *op. cit.*, p. 93.

[258] G. L. Bell, "Notes on a journey through Cilicia . . ." in: *Revue Archéologique*, sér. IV, VI (1906), p. 409; J. Strzygowski, *Kleinasien . . .* , Leipsic, 1906, p. 54, Figs. 40, 41; J. T. Bent, "A Journey in Cilicia Tracheia" in: *Journal of Hellenic Studies*, XII (1891), p. 206 ff.

[259] G. Weber, "Basilika und Baptisterium in Gül-bagtsché", in: *Byzantinische Zeitschrift*, X (1901), p. 570 ff.

[260] Herzfeld and Guyer, *op. cit.*, p. 112 ff., date the edifice at the end of the VI century; Kautzsch, *op. cit.*, p. 89, dates it in the first half of the VI century.

[261] W. R. Ramsay and G. L. Bell, *The Thousand and one churches*, London, 1909, p. 209 ff.

[262] H. H. Jewell and F. W. Hasluck, *Our Lady of the Hundred Gates*, London, 1920.

[263] J. W. Crowfoot, *Gerasa*, New Haven, 1938, p. 182 f., p. 249 ff.

[264] Gsell, *op. cit.*, p. 170 ff.; the church is not dated.

[265] U. Monneret de Villard, *Les couvents de Sohag*, Milau, 1925 f., passim.

In groundplan this layout resembles very much the center bay and the wings of a tripartite transept, and sometimes it is quite impossible indeed to make out whether or not it is actually such a tripartite transept. As a matter of fact, it has been frequently interpreted as a transept, for example in the "transept basilica" at Korykos where the church has even been named from this pattern. Yet this interpretation can hardly be maintained in every single case; first, what resembles the wing of a transept is, in these churches, a small square bay which in its interior could hardly have reached the height of the center bay, that is, of the forechoir and of the nave. If it had, it would have created in the interior the impression of an extremely narrow and tall shaft-like structure. One would assume rather that this bay was either lower than the center bay in its exterior as well as in its interior, thus resulting in a dwarf transept, or else that it was subdivided into an upper and a lower story, thus being lower in the interior while corresponding to the height of the center bay in its exterior. Indeed the remnants of such a division into two stories are actually preserved at Kanytelideis where the photographs and drawings published, clearly show lower and upper openings leading from this bay at the end of the aisles into the aisle as well as into the forechoir (Fig. 18).

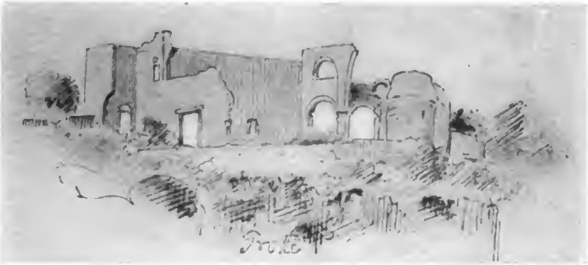


FIG. 18. Kanytelideis, Basilica IV, drawing (J. Strzygowski, *Kleinasien*, 1904, Fig. 40).

---

[266] G. Ph. Stevens, J. M. Paton, a.o., *The Erechtheum*, Cambridge, Mass., 1927, p. 492 ff.

Whether such an arrangement with upper stories existed in any of the other buildings of the group can no longer be ascertained, since in none of them enough is preserved to allow for a definite statement. But the arrangement of two-storied transept pastophories at Kanytelideis opens a wide perspective.

On the one hand, this solution seems to be related to the basilica of Kodja Kalessi in Cilicia, which is usually considered a typological forerunner of the Byzantine domed basilicas, such as H. Irene at Constantinople, although chronologically it is now dated later.[267] The eastern part of Kodja Kalessi with a square bay in front of the apse, surmounted by a towerlike structure, probably with a wooden roof, not a stone dome, and flanked by two bays with galleries on the upper stories, the emphasis given to the center bay and the fact that this center bay is clearly marked off as a forechoir, all point towards a close relationship between Kanytelideis and Kodja Kalessi. Since neither of these two structures can be dated with any degree of precision, the exact relation cannot be clarified at present. Both may be representatives of types which had existed in the Near East since the early V century.[268]

While this connection is important for the history of architecture in the Near East, the interrelation between the type represented by Kanytelideis and the type represented by S. Pietro in Vincoli and by the related group of Greek churches becomes decisive for our problem.[269] Evidently the transept pastophories have grown in size so as to become genuine wings of a tripartite transept; yet these wings are clearly marked off from the center bay by the arrangement of an arcade. The existence of an upper wall with openings, which can be established at least for the first project of S. Pietro in Vincoli, viewed in the light of this excursion into the problem of the pastophories, appears to be the rudiment of an upper story of pastophories of the type preserved at Kanytelideis. In the successive stages of S. Pietro in Vincoli, in the first project with upper walls and openings, and in the following stage where these upper walls with their openings were omitted, and possibly also the arcades between center bay and wings on the ground floor were discarded, as they certainly are at Nikopolis A and in the related Greek basilicas of different types, the tripartite transept was created.

---

[267] C. Headlam, "Ecclesiastical sites in Isauria", in: *Journal of Hellenic Studies*, Supplementary papers, 2 (1891), p. 9 ff., dated the building, which he was the first to publish, in the early V century; Kautzsch, *op. cit.*, p. 96, dates it now after the middle of the VI century, basing this date on the capitals.

[268] The church of the Crocefisso at Spoleto in Italy seems to be related to the type of Kodja Kalessi; while it has no galleries over the bays at the end of the aisles, these bays, as well as the center bay with its domed tower, are clearly similar to Kodja Kalessi. Since Spoleto seems to have been built about 400, this would strengthen the above mentioned hypothesis. See A. Salmi, "La basilica di San Salvatore . . ." in: *Dedalo,* II (1921/22), p. 628 ff.

[269] See also Soteriu, *Ephemeris,* 1929, p. 198.

It should be emphasized that the different types discussed, pastophories, intermediate forms, and wings of the tripartite transept, need not have followed each other always in a clear chronological sequence. What occurred was evidently that from the V century.on, a general attempt was made to enlarge the pastophories and to link them more closely to the aisles and to the forechoir. The solutions found were evidently different in different regions, and what seems to have resulted logically from an earlier step may actually have been an independent local solution. The different types of pastophories with their different entrances, the transept pastophories with and without upper stories, the forerunners of the domed basilicas, and the genuine tripartite transept apparently evolved side by side. It is very possible that this last pattern, the tripartite transept, in its different stages, developed either in Greece or at S. Pietro in Vincoli in Rome under the influence of the continuous transept which had been so common in the Roman basilicas of the IV century. The Roman prototype of St. Peter's or S. Paolo f.l.m. may easily have caused the transformation of a pastophory into the wing of a tripartite transept, as it caused at the same time in the Near Eastern Coastlands the transformation of the cross-shaped memoria into the churches with cross transepts.

JOHN BOWRING AND THE POETRY OF THE SLAVS

ARTHUR PRUDDEN COLEMAN

Department of East European Languages,
Columbia University

(*Read Nov. 23, 1940*)

ABSTRACT

Sir John Bowring (1792–1872), most commonly remembered as the friend and literary executor of Jeremy Bentham, was "one of the few partisans of the cultivation of foreign literatures in England", according to R. P. Gillies, who knew him well. A close examination of Bowring's contributions to the study of the poetry of the Slavs confirms the truth of Gillies' statement. Besides his two volumes of *Specimens of the Russian Poets*, Bowring published one each of translations from Polish, Serbian and Czech. At least three of these won quick popularity, and the Russian volume was widely circulated in the United States, where selections from it were generously reprinted in readers for the Fifth and Sixth Grades. Bowring's most popular translation was that of the poem *God* by the Russian courtier Derzhavin; this was reprinted in two generations of American readers.

Bowring's interest in Slavonic poetry arose as a natural concomitant of the Utilitarian philosophy of his master Bentham. He desired earnestly to do his part in ushering in Bentham's ideal world, wherein "the greatest possible good" should be guaranteed to "the greatest number of people". In harmony with this purpose and contributory to its fulfillment, Bowring early conceived the desire to become a peddler of poetry; while his countrymen denounced this race or that as barbarian, he, John Bowring, would publicize the poetry of even the least likely, so that none could with justice be called uncultivated. "I have never left the ark of my country", he said, "but with the wish to return to it bearing fresh olive branches of peace and fresh garlands of poetry."

Living in the very "springtime of the Slavs", the early 19th century, Bowring carried on a lively correspondence with eminent Slavonic scholars. While his method was that of a *dilettante* and evoked many quips and sallies from his countrymen, one must admit, nevertheless, with the *Edinburgh Review*, that "He circulates the produce of thought, varies our intellectual banquets, teaches us that some accession to our store may be derived from those quarters which we had regarded as the most sterile and unpromising and thus adds another link to the chain of social and friendly feeling which should bind man to his fellows."

The present study is the first serious appraisal of Bowring as he appeared to his contemporaries, both English and Slavonic.

I

In 1843, twenty years after the publication by John Bowring (1792–1872) of the second volume of his *Specimens of the Russian Poets*, the first in a famous series of translations from the Slavonic,[1] Thomas B. Shaw accused Bowring of "gross unfaithfulness" to the original text of the poems he claimed to translate. Shaw branded Bowring's carelessness "the less excusable, as the verses in question present nothing in style, subject, or diction, which should offer the smallest difficulty to a translator.[2]" What Shaw, who knew Russian thoroughly himself, meant to imply was that Bowring's knowledge of the Russian language was inadequate for a translator.

---

[1] *Specimens of the Russian Poets*, Vol. I, London, 1821, Vol. II, 1823; *Servian Popular Poetry*, 1827; *Specimens of the Polish Poets*, 1827; *Cheskian Anthology*, 1832.

Shaw might have pointed out, had he wished, an inadequacy no less grave in Bowring's knowledge of Polish, Serbian and Czech. Yet this remarkable man published, besides the two volumes of Russian *Specimens*, one each of Polish, Serbian and Czech. Of these at least three won quick popularity. The Russian volume was widely circulated in America and its selections reprinted in American readers,[3] and all five anthologies sold well enough to net their compiler a financial profit.

How did Bowring do it? And how did his contemporaries regard him as a translator of Slavonic poetry?

John Bowring's ruling purpose as he reached maturity was to make effective the Utilitarian philosophy of his master, Jeremy Bentham. He desired earnestly to do his part in ushering in Bentham's ideal world, wherein "the greatest possible good" should be guaranteed to "the greatest number of people". In harmony with this purpose, and contributory to its fulfillment, Bowring early conceived the desire to become a peddler of poetry; while his countrymen denounced this race or that as barbarian, he, John Bowring, would publicize the poetry of even the least likely, so that none could with justice continue to be called uncultivated. "I have never left the ark of my country", he wrote, "but with the wish to return to it bearing fresh olive branches of peace and fresh garlands of poetry.[4]"

---

[2] Thomas B. Shaw, Introduction to a translation of *Amalát Bek*, *Blackwood's Magazine*, LIII, Mar. 1843, p. 283. Shaw, a Cambridge University M. A., was in 1843 Adjunct Professor of English Literature in the Imperial Lyceum of Tsarskoë Selo.

[3] The most popular volumes were the two Russian and the Serbian. Selections from the Russian volume were reprinted almost at once in the Rev. John Pierpont's *American First Class Book in Reading and Recitation*, Boston, June, 1823, a book which was immediately adopted for use in the public schools of Boston. The following Bowring translations were used in this work:

1. Lomonosov: "The Lord and the Judge", p. 93.
2. Khemnitzer: "The House-Builder", pp. 93–94.
3. Khemnitzer: "The Rich Man and the Poor Man", pp. 378–379.
4. Dmitriev: "Lines written during a Thunderstorm", p. 96.
5. Derzhavin: "The Waterfall", pp. 366–367.
6. Derzhavin: "God", pp. 475–478.
7. Karamzin: "The Churchyard", pp. 377–378.
8. Krylov: "The Ass and the Nightingale", pp. 467–468.
9. Bobrov: "Address to the Deity", pp. 469–471.

The two most popular Bowring translations were those of Derzhavin's "God" and Karamzin's "The Churchyard". These were often reprinted in readers for at least a generation. A few such reprintings are the following:

Of "God":

1. Arethusa Hall, *Literary Reader*, Boston, 1851, pp. 406–408.
2. Epes Sargent, *Standard Fifth Reader*, Boston, 1856, pp. 153–155.
3. G. S. Hilliard, Sixth Reader, Boston, 1865, pp. 372–373.
4. Marcus Willson, *United States Series*, New York, 1872, p. 364 [in part].
5. C. W. Sanders, *Fifth Reader*, New York, 1868, pp. 162–165.

Of "The Churchyard":

1. Lucius Osgood, *Progressive Fifth Reader*, Pittsburgh, Pa., 1858, pp. 210–211.
2. G. S. Hilliard, *op. cit.*, pp. 372–373.
3. Marcus Willson, *op. cit.*, pp. 107–108.

It was during the Slavonic renaissance of the early XIXth century, in the very "springtime of the Slavs", that Bowring left "the ark of his country" for the first time and it was this excursion that gave him his start as a peddler of verse.

Visiting Russia as a young merchant in the winter of 1819–1820, Bowring found the Slavonic world buoyant with poetic vitality and eager to have its literary wares peddled abroad, especially in England. The fact that he knew only a smattering of Russian picked up from a German text,[5] did not stand in Bowring's way. He filled his pack with paraphrased versions in English and German of Russian poems which his friends in St. Petersburg selected for him and hastened home to publish its contents, turned by himself into English verse, under the title *Specimens of the Russian Poets.*

Hardly had this volume, which Bowring himself admitted was in the nature of a "taster", come from the press than the *London Magazine* gave it a fine send-off. This was to be expected, since Bowring was one of the *London Magazine's* friends politically and a frequent contributor to its pages. The reviewer found in the *Specimens* much to praise, above all Batiushkov's "To My Penates", a pleasing rendition which years later W. H. Leeds called a translation "at once accurate and spirited".[6] "Upon the whole", the *London Magazine* concluded, "we consider this volume as one of the most agreeable and interesting that has come before us for some time past. It was put into our hands quite unexpectedly and very late in the month; but we have proceeded to notice it without delay, both on account of the public, who will be anxious to know the character of a work on so novel a subject and that the translator may not remain in doubt as to its probable reception."[7]

---

[4] A sentiment reiterated in different words in every article Bowring wrote.

[5] The German text used by Bowring was Tappe's *Neue russische Sprachlehre für Deutsche*, St. P., 1820.

[6] *Foreign Quarterly Review*, IX, Jan. 1832, p. 220. W. H. Leeds (1786–1866) wrote critical articles on architecture and was a frequent contributor to the *Foreign Quarterly.*

The amazing interest which the public did show in the little volume arose, as the *Edinburgh Review* decided[8] ten years afterward, from wonder, for a volume from Russia which betrayed "no marks of its barbaric origin" was bound to attract attention in 1821 through its very failure to conform to the English notion of what should properly come out of Russia. Across the Atlantic, in Boston, where the book was quickly issued in an American edition, the same emotion greeted it: ". . . we were almost as much taken by surprise", admitted the Reverend Mr. Peabody in the *North American Review*, "as by the ode which Major Denham brought from the court of his colored majesty of Bornu."[9]

In Boston the Reverend John Pierpont used nine of Bowring's translations in his *American First Class Book in Reading and Recitation* (1823). Thus in America as in England Bowring created a taste, as Peabody said,[9] which he was ready and even eager to gratify; he quickly set to work on a second Russian collection.

It had been the able assistance of Friedrich Adelung, the distinguished nephew of Johann Adelung of *Mithridates*, that enabled John Bowring to undertake a volume of Russian verse and to carry it through in spite of his poor linguistic equipment. Adelung, who stood high in imperial favor in Russia through his position as tutor to the Grand Dukes, received Bowring kindly in St. Petersburg and either prepared himself or assembled from the hands of others translations of Russian verse into the languages Bowring understood. Now when Adelung learned that the second volume of *Specimens*, which he had planned with Bowring while the latter was still in Petersburg, was to become a reality, he continued to help. He not only advised Bowring at length of orthographic errors in the first volume that should not be repeated in the second but he kept him supplied also with new books like Nicholas Grech's *Short Survey of Russian Literature* (1822) which he called an "excellent manual".[10] From time to time also he sent additional Russian verses, one of these Zhukovsky's "Epistle to Tsar Alexander" in a German version made by a Mr. Rauspach, a professor in the University of St. Petersburg. Adelung tried also to get Krylov to make English versions for him of his Fables, so that Bowring could include more of these in his second volume. In return for this expert guidance on the two volumes as well as for Adelung's handling in Russia the sale of the *Specimens* and their distribution among scholars, Bowring reciprocated by sending his friend the works of Bentham and by supplying him with the English and American reviews of Adelung's own *Uebersicht aller bekannten Sprachen und ihrer Dialekte.*

---

[7] *Op. cit.*, III, Mar. 1821, p. 321.

[8] "Dr. Bowring's Poetical Translations", *Edin. Rev.*, LII, Jan. 1831, p. 323.

[9] *Op. cit.*, XXVI, Jan. 1828, p. 146. Peabody refers here to the sensation of 1824, Dixon Denham's findings in Central Africa.

[10] Frequent references will be made, from this point on, to letters exchanged between Bowring and his Slavonic friends and collaborators. These letters may be found in *Kralovská česka společnost nauk*, Věstnik, 1904 and 1912, under the following titles: "Korrespondence Johna Bowrings do Čech", assembled by Robert Beer, and "Listy psané J. Bowringovi ve věcech české a slovanské literatury", F. Chudoba, editor. Since

Unfortunately Bowring was in no position to make use of most of the help Adelung offered him on the second volume of *Specimens.* Much of it arrived while he was in Boulogne Prison in Calais where the French government detained him for several weeks in 1822 on account of his alleged anti-Bourbon activities. Here Bowring finished the second book in October, having passed the long hours of his imprisonment pleasantly, as he said, in reflection on Slavonic song and on his friends in Russia who relied on him to make it known abroad.

Published early in 1823, the Russian *Specimens,* Volume II, had to wait until November, 1827, four years, for a review in an important journal. Then, however, it received what most of Bowring's translations from the Slavonic never received, serious criticism by one who understood thoroughly both the Slavonic tongue and the English. The man who performed this service was Victor Smirnov, a secretary in the Russian Embassy and son of the Embassy's chaplain. In order to weigh the merit of Bowring's translations from the Russian, Smirnov, together with John G. Cochrane, editor of the *Foreign Quarterly Review,* compared his English version of Zhukovsky's "Svetlana" with the French version of Dupré de St. Maure's *Anthologie Russe* (Paris, 1823) which appeared almost simultaneously with Bowring's own. Side by side Smirnov and Cochrane placed the two renderings, together with a literal English version which Smirnov himself made. Their conclusion after examining the three was that "if the English adheres more closely to the original in the *measure,* and exhibits greater poetical condensation, the French one is not

---

the substance of these letters is woven into the very fabric of this article to such an extent that complete annotation would overload the text, no specific citations from these two sources will be made.

inferior to it in *literal* fidelity, and by its amplification tells the story more clearly and distinctly", adding, finally, "Beauties there are in the original which we think have escaped the grasp of both translators".[11]

II

As the second volume of Russian *Specimens* came from the press, there appeared in Edinburgh a work by the Pole Christian Lach Szyrma called *Letters, Literary and Political, on Poland* (1823), in which Bowring's first Russian collection was recommended as an "excellent" survey.[12] No sooner had Bowring read this than he addressed himself to the author of the *Letters*, who was already known to him from articles he had published in *Blackwood's*,[13] proposing that they two collaborate on an anthology of Polish verse. Bowring was ready now to act as promoter of Polish literature, a role he had carried off with so much success in Russian. Being ready to do this ten years before London swarmed with refugees from Poland, when help in such an enterprise was easy to obtain, he was casting about for a collaborator. From the broad-minded, sympathetic and well-informed tone of his *Letters* Szyrma seemed to Bowring the very man he was looking for; here, perhaps, would prove to be his Polish Adelung!

Flattered by Bowring's notice and by his praise of the *Letters*, Szyrma replied at once, tackling in his very first letter the details of Bowring's proposal. Inquiring first whether he planned to have the Polish volume follow the pattern of the Russian, Szyrma asked Bowring whether he knew enough Polish himself for such an undertaking and what equipment he already had. He even offered to send to Warsaw for books on his own account and to make the selection of these himself so as to spare Bowring the embarrassment of not knowing what to order. In Edinburgh, he added, there was a small collection of Polish books in the Signet Library, deposited there by Count Andrew Zamojski; these Szyrma promised to secure, if possible, for Bowring's use.

Szyrma's response was welcome indeed; Bowring replied at once, putting question after question to the Pole and suggesting at the same time the lines along which he proposed to build the Polish anthology. Soon Szyrma began to send actual material: Onufry Kopczyński's standard Grammar of the Polish Language; a Bandtkie French-German-Polish Dictionary and several volumes of poetry from the famous series published in Warsaw from 1803 on under the editorship of Tadeusz Mostowski as *Wybór pisarzów polskich* (*Anthology of Polish Writers*). Szyrma sent also Niemcewicz's famous *Śpiewy historyczne* (*Songs from History*) and probably his *Bajki i powieści* (*Fables and Tales*), besides scattered numbers of the contemporary literary journal *Pamiętnik warszawski* (*Warsaw Memoir*). Bowring's questions Szyrma did not stop now to answer, preferring to get his specimen translations off promptly with the books on the London packet. He advised Bowring to use Herder if he needed help with his Lithuanian selections, a suggestion Bowring followed, using two from Brodziński's collection in *Pamiętnik warszawski* which Herder had long before included in his *Stimmen der Völker in Liedern* (1778–9).

---

[11] "Russian Literature", *FQR*, I, Nov. 1827, p. 600.

[12] *Op. cit.*, p. 80.

[13] "Slavonic Traditional Poetry", *op. cit.*, X, Sept. 1821, pp. 145–151; "English Literature in Poland", *op. cit.*, XI, Mar. 1822, p. 329; Review of Józef Wawrzyniec Krasiński's *Guide du Voyageur en Pologne et dans la République de Cracovie*, Warsaw, 1820, *op. cit.*, June, 1822, pp. 650–657.

In July, 1823, Bowring came to the point in his work on the Polish anthology where, if the book was to go on, he had to have historical and critical material that he could read himself. Fragmentary articles he found, on Szyrma's suggestion, in back numbers of the *Göttingische gelehrte Anzeigen*, but these contained nothing that could take the place of what Bowring needed at this moment and did not have: a mellow, comprehensive knowledge of Poland's literary history.

The inadequacy of Bowring's Polish and of his method as a translator stand out sharply in his letters to and from Szyrma. Bowring was baffled, in the first place, by Polish diminutives. Secondly, he neither felt nor made any particular effort to feel the imagery of the original. Thus in Kochanowski's seventh "Lament" he could fly so far from the poet's concept as to represent the grief-stricken father in the act of gathering hay (!) for a pillow to place beneath his child's cold head, when the imagery of the Polish ("ziemie bryłeczkę") was the symbolic, tradition-hallowed Slavonic custom of collecting handfuls of earth. Even when Szyrma pointed out this glaring travesty Bowring did not correct it; all he did was change the "balmy hay" of his first version to "some odorous hay". In his rendering of the song "To Sleep" he showed himself equally unconcerned with the sense of the original when he refused to change the last lines in accordance with Szyrma's suggestion. Bowring was satisfied, obviously, if his verses adhered to the original in form and measure, whether they captured the imagery of the Slavonic or not.

Yet Szyrma was pleased, on the whole, with Bowring's versified renditions of the paraphrases he sent him. He called Bowring's "duma" of Michał Gliński so good, after reading it over several times, each time "with increasing pleasure", that he could suggest no improvements, and he praised as "beautiful" and "masterly" Bowring's version of the exquisite song "To Sleep".

In the summer of 1823 Szyrma began trying to shift the responsibility he had incurred for the volume of Polish poetry to the shoulders of someone more competent. The man he selected for the difficult honor was Kazimierz Brodziński, whom Szyrma himself called, in an article in *Blackwood's*, "the most distinguished young poet in Poland".[14] Though only in his early thirties, Brodziński was already famous as an interpreter of western ideas in Poland because of his translations from French, German and English. Not long previously his translation of Scott's *Lay of the Last Minstrel* had established him as an outstanding Anglicist along with Władysław Ostrowski and Karol Sienkiewicz. Brodziński was currently very popular in Warsaw as a consequence of his able lectures in the University, and he had wide influence, moreover, as an editor of *Pamiętnik warszawski*. He seemed the very man for Bowring.

After repeated solicitation from Szyrma and a letter from Bowring himself, Brodziński showed active interest in the proposed anthology. In October, 1823, he thanked Bowring for his desire to serve the Polish race and promised his help. He warned the over-optimistic Englishman, however, that he would require much more information than he had yet received before he could go ahead. To Bowring's request for specimens of Polish verse translated into some language he could read, Brodziński replied with the question "Into what language?", adding that he would prefer German, since his own English was inadequate and French entirely unsatisfactory as an intermediary tongue. He asked Bowring how long the summary of Polish literary history he was to write should be and assured him of his intention to prepare translations of poets like Woronicz and Kniaźnin, hitherto neglected abroad.

Here was an auspicious beginning indeed. And the more auspicious because Szyrma too seemed not to have abandoned helpful activities even though he had left Britain to return to his native land. From Paris in January, 1824, he wrote to thank Bowring, and Mrs. Bowring too, for receiving him so pleasantly during his stay in London, as well as to report that he had looked vainly in Paris for a copy of La Garde's translation of Trembecki's "Zofjówka" (1815). Szyrma's final word as he left Paris was a friendly and an optimistic one, remarking that his own *Letters* had been favorably noticed by the Polish press, adding "you may expect still more as a foreigner", and "I will remember you in Poland."

---

[14] See the second article mentioned in note 13, *Blackwood's*, XI, p. 329.

When Szyrma was heard from again it was more than six months later and then with an apology for the silence of his friend Brodziński and for his own. The poet had gone, he said, to Italy for a long holiday and he himself had been too busy getting settled in Warsaw to think of anthologies. As a peace offering he sent Bowring with this letter a ballad of his own called "A Child at his Mother's Knee", which Bowring did not use.

While he was waiting for Brodziński's help to materialize, Bowring put out a couple of feelers to gauge the popular appetite for Polish verse. These were two Polish songs, one a traditional verse, "The Three Fountains", the other a poem from Zimorowicz beginning "I saw thee from my casement high". Both of these, done with Szyrma's help, appeared in the *London Magazine*, February, 1824.

III

Bowring now found himself a very busy man. His linguistic interests continued varied and absorbing; at the moment he was not only finishing his Batavian and Spanish anthologies and a collection from the German called *Matins and Vespers*, but he was corresponding at the same time about a proposed Finnish collection and a Swedish. He had, moreover, just been appointed editor of Jeremy Bentham's new *Westminster Review*, so that for months he was deeply involved in the dissensions that marked the launching of that journal. His plain editorial duties were likewise by no means light. Having been assigned to the editorship of the literary department of the *Review* by Bentham largely because of his wide foreign friendships, Bowring was expected to contribute to it frequent articles on world thought. In the *Westminster's* first number, therefore, January, 1824, appeared his first article: "Politics and Literature of Russia".

The article is a demonstration *par excellence* of Bowring's technique as reviewer and critic. Ostensibly it reviewed two works: M. de Pradt's *Parallèle de la Puissance Anglaise et Russe relativement à l'Europe* (Paris, 1823) and a sketch of Russian literature by Alexander Bestuzhev which appeared in *Polyarnaya Zvyezda* for 1823. The portion that reviewed the latter was impressive indeed with its battalions of Russian proper names. It was, however, as *Fraser's* reviewer pointed out fifteen years later,[15] nothing more than an English rendering of a German version of Bestuzhev's article that appeared in Vol. IX of Oldekop's *St. Peterburgische Zeitschrift.*[16] Bowring must have read the German review hastily and have sent his own article to press without checking his translation, for out of "*Der blinde Eros* beweist dass Siberiens Winterfröste seine (Sumarokov's) scherzhafte Phantasie nicht erkalteten",[17] he made "and a Siberian bard, the blind Eros, published a volume of jocose poetry"! As Fraser's put it, "A more singular literary *quid pro quo* is not, perhaps, upon record."[15]

Through applying and perfecting the technical method exhibited in the article on Russian literature, Bowring became, as even his bitter critic John Neal had to admit, a kind of High Priest of Critics whenever the poetry of the Slavs was under consideration. Two factors besides his skill in using German sources contributed to his virtual episcopacy: the salesman's self-confidence that was native to him and the flattery accorded him by all his Slavonic collaborators except, as we shall see, the Slovene scholar Yerney Kopitar.

In his capacity of critic Bowring placed fidelity to the measure and verse-form of the original first in importance. Failure on the part of a translator to observe this particular kind of fidelity provoked his critical censure. Thus, in reviewing for the July, 1825, *Westminster* a polyglot edition of Krylov's *Fables*,[18] he waxed caustic over the process by which these found their way into French and Italian verse. Count Orlov, who compiled the volume, took some of the fables from the Russian edition of 1819 and turned them into French prose; these rough paraphrased versions he then handed to French and Italian poets and they, in turn, wrote in their best poetic manner "imitations" of the Russian verses, trying in these imitations to retain the essential meaning of the original. Count Orlov spoke of this process with considerable pride, claiming it had the advantage of releasing the poet "from the trammels of the original", meaning, of course, the trammels of form and measure. Bowring expressed nothing but condemnation for such a system and, to illustrate the more faithful manner in which he would do it, provided three Fables in an English version of his own, made presumably from Krylov's rough translations sent by Adelung too late for inclusion in the second *Specimens.*

---

[15] *Fraser's Magazine*, XIX, Feb. 1839, footnote to pp. 159–160.

[16] *Op. cit.*, IX, pp. 129–156. "Blick auf die alte und neue Literatur in Russland" von A. Bestushew, nebst einem Nachtrage des Herausgebers.

[17] *Op. cit.*, p. 139.

[18] *Fables russes, tirées du recueil de M. Kriloff, et imitées en vers francais et italiens par divers auteurs*, Paris, 1825, 2 vols.

IV

Bowring's editorial duties and his private business obligations were heavy now (1825), but they did not prevent him from prosecuting his plan of bringing to English attention all Slavonic literature. "As I have in some respects a sort of vague feeling that I have the Slavonians under my charge, whenever I can do them honor, I shall from time to time scribble such matters as are likely to extend the good feelings of my countrymen", he wrote to one of his Slavic friends. Now, though his Polish anthology had reached a stalemate, a new and more than usually interesting Slavonic personality came into Bowring's orbit.

For years everyone who followed continental literature by reading the German literary journals had been aware of the significant work being accomplished by the Serbian scholar Vuk Stefanovich Karajich. Since 1814 Karajich had been publishing collections of Serbian song and story and now in 1825 comments were appearing not only on his own most recent work, *Srpske Narodne Pyesme* (*Serbian National Songs*), the first three volumes of which had appeared in 1823 (Leipzig), but also on a German rendering of some of these by Therese Albertine von Jakob (known as "Talvj"), in the first volume of her *Volkslieder der Serben* (Halle, 1825). Both Karajich's collection and Talvj's translations fell quickly into Bowring's hands and soon he was no less enthusiastic about Vuk Karajich and Serbian poetry than he had been earlier about Krylov and Derzhavin and Russian verse and later about Niemcewicz and his heroic ballads.

Talvj's *Volkslieder* was exactly what Bowring needed at this juncture. Having this before him he could follow line by line, with the aid of Karajich's Serb-German dictionary (1818), the original Serbian of the *Narodne Pyesme.* As a matter of fact Bowring probably did very little of this hard plodding work. For the most part he simply paraphrased Talvj's German verses. Some years later when the American poet James Gates Percival discovered Bowring's outright filching of Talvj's notes without crediting her with their origin, he was angry and declared the man showed "the cloven hoof".[19] Talvj herself did not resent Bowring's use of her work, though she knew he did it, and on more than one occasion recommended his verses highly.[20]

The first fruit of Bowring's interest in the poetry of the Serbians was not long in appearing; it came in the form of a sixteen page article in the *Westminster Review* for July, 1826, on "The Popular Poetry of the Serbians". This was frankly an *hors d'oeuvre;* the *pièce de résistance,* a full volume of Serbian verse, was already cooking in his mind.

The necessity of thanking Paul Joseph Šafarik for a copy someone had sent him of *Geschichte der slawischen Sprache* gave Bowring now an opportunity to make inquiries of a personal nature about Karajich. From Šafarik he was able to learn details of the early life of this lame Serbian folklorist and something also of his present situation—how he was living in Vienna on a pension granted by the Tsar Alexander. From him, too, Bowring learned that the best critical material on Karajich was Kopitar's review of the *Pyesme* in the *Wiener Jahrbücher der Literatur* for 1825. Šafarik called Bowring's attention also to Eugene Weselý's[21] German version of some of Karajich's songs, calling it, however, "far inferior to Talvj".

No sooner had Bowring received Šafarik's reply than he wrote personally to Karajich, for whom apparently he quickly conceived a warm sympathy, a letter in which he laid out his plans for a Serbian anthology. When Karajich replied, he was on the point of leaving Vienna for Serbia on a new song-collecting expedition under the patronage of the ruling prince, Milosh Obrenovich. Besides thanking Bowring for his desire to honor "the forgotten Serbian nation" and asking Bowring for help on his own projects, Karajich sent books that he believed might prove useful, among them the volume Šafarik had mentioned, Weselý's *Serbische Hochslieder*. Most important of all, Karajich referred Bowring to the great Slavonic scholar Yerney Kopitar as the person above all others competent to guide him.

---

[19] See letter of Percival to Talvj, Feb. 26, 1836, in H. R. Warfel, *Unpublished Biographic Study of James Gates Percival,* pp. 444–453, on file in the Library of Yale University, New Haven, Conn.

[20] Talvj, *American Biblical Repository,* IV, April 1834, pp. 328–413, and July 1834, pp. 417–531.

[21] *Serbische Hochslieder*, Buda, 1826.

The Serbian volume was too far along and Bowring too impatient for publication to wait for advice from Kopitar. Relying for his introductory material on Šafarik and on German authorities suggested by Šafarik; on Grimm's introduction to Karajich's *Kleine serbische Grammatik* and on his friend Adelung's *Uebersicht aller bekannten Sprachen* for the German text of his verses; and for his notes on Talvj's *Volkslieder* and Weselý's work, Bowring put the Serbian anthology together and published it early in 1827 as *Servian Popular Poetry*.

The first journal to notice the book was the *London Magazine*. In April, soon after its publication, the *London's* reviewer had this to say about it, "Mr. Bowring's felicity in the difficult art of translating poetry is well known to all lovers of it. Together with an almost marvellous knowledge of the dialects of Europe, he possesses a ready tact in seizing the tone and character of his subject. His poetical sympathies are so warm and prompt that it would be impossible to place him in the midst of any class of ideas or feelings where he would not almost instantaneously adapt himself to the hue and colour of the imaginative circumstances about him. His command over the stubborn materials of his own language is very considerable, which more especially qualifies him for the task he has voluntarily chosen of throwing his translations into the measures of the original. Of the fidelity of his Servian versions we are wholly unable to judge; internal evidence would lead us to suppose it was close." [22]

Bowring had the distinction of being the only man of his time in Britain to translate Russian poetry; in Serbian he had no such distinction. Scott's interest in the Serbian ballad, a by-product of his Goethe studies, was of some years' standing and he too dabbled at translating bits of it from the German version. Now, moreover, just as Bowring's Serbian anthology was issuing from the press, other translations of Serbian poetry were appearing in the *Quarterly Review* [23] which even the highly pro-Bowring *London Magazine* had to admit would, because of their "poetic embellishment and some adaptation" be more likely to "attract the attention of the mere English reader".[24]

---

[22] *Op. cit.*, XVII, Apr. 1827, p. 569.

[23] "Translations from the Servian Minstrelsy", *op. cit.*, XXXV, pp. 66–86.

A distinction of a different kind Bowring still could claim: he alone could make his readers feel that the poet whose verses he translated was his close friend. As Jared Sparks put it in the *North American Review* (October, 1827), "With the minstrels of Russia, Poland, Servia and other countries he is apparently as much at ease as if they had been his intimates from infancy under his native skies, and he enters with readiness into their local associations, feelings and attachments, looking out upon the broad compass of nature, tracing its varied forms, and recognising the manners and social habits of different nations and ages, with a familiarity that would seem to betray an intimate companionship." [25]

Bowring gave his readers this impression partly by such warm, personal dedicatory verses as those to Karajich which introduced his Serbian anthology:

My friend! it is thou, it is thou
  Who hast ushered these gems into day;
'Tis my pride and my privilege now
  To honour—I fain would repay
Thy toils, and would bind round thy brow
  The laurels that grow o'er thy lay.

and which end with the lines:

Thy tenement is but of clay;
  Thou art frailer than most of us be:
Yet a sunshine has lighted thy way,
  Whose effluence is sunshine to me:—
And 'tis sweet o'er thy Servia to stray,
  And to listen, pale minstrel, to thee.

As the *London Magazine* put it:

> Through the whole of Mr. Bowring's writing this warm and generous sympathy with foreign and distant individuals, whose tie to him is solely that of kindred labours, is highly characteristic. The same facile and generous sympathy, not only with persons but with their feelings, their habits and their language, renders Mr. Bowring not only one of the most amiable of men but one of the ablest and readiest transfusers of the spirit of national poetry.[26]

---

[24] *London Magazine*, *op. cit.*, p. 574.

[25] *North American Review*, XXV, p. 355.

[26] *London Magazine*, XVII, April 1827, pp. 582–583.

A saltier critic, not taken in by Mr. Bowring's charm of manner and his talent for friendship, found something cloying in all this warmth and intimacy and adaptability. To all such men as Bowring he addressed a fable which began,

Writers, to you my tale's addressed,
Who with most fertile pens are blest;
Yet gifted with but small discretion,
Would feast the public to repletion;
Force down their throats book after book,
Cramming them till they fairly choke;
And of your prose, or verse to tuneful,
Insist on giving us another spoonful.[27]

This reaction came, however, only with the advent of *Fraser's Magazine* and long after Bowring had passed his zenith as a translator.

The only important journal besides the *London Magazine* to notice the Serbian volume promptly was London's most aggressive literary magazine, Colburn's *New Monthly*. Like the *London*, it was of the Liberal faith and friendly to Bowring. The reviewer on this occasion took exception rather strongly to Bowring's guiding principle as a translator of verse, his faithful adherence to the measure of the verse. It would have been better, he believed, to use the English ballad measure, with its peculiar rhyme and its charm for the English ear, than to have left the verses in the characteristically Serbian, unrhymed state. Certain terms which Bowring's fidelity to the letter caused him to use were also objectionable to this reviewer—his use, for example, of the adjective *dexter* to describe a bird's wing. "With the exception, however, of a few trivial mistakes of a similar kind . . . we can scarcely speak too highly of the general merits of the specimens before us. Many of them lay claim to real poetry and passion—to bursts of simple and heartfelt pathos, like unpremeditated and improvised effusions, and notwithstanding the difficulties imposed by Mr. Bowring's system of perfect fidelity, they are often as beautifully as they are faithfully rendered." [28]

On the whole, the Serbian translations, though never copied in textbooks to the extent the Russian *Specimens* were copied and never reprinted as some of the Russian poems were, received more widespread approval and more sincere admiration than Bowring's other collections. This was due in the first place to the exotic nature of the songs themselves and secondly to the inherent merit of Bowring's rendition of the poems. Here, as Jared Sparks said, he did make the verses "live and breathe" with all their "original force and peculiarity".[29] The fire Bowring caught from Goethe's enthusiasm for Serbian folk verse and from Talvj's splendid German translations he was often successful in transmitting to the English reader. Prosper Mérimée's hoaxing in *La Guzla* attracted popular attention to Serbian verse just about the time Bowring published his volume and the publicity which attended the reception of this astonishing work contributed in no small degree to the success of Bowring's little *Servian Popular Poetry*.

---

[27] "Russian Fabulists with Specimens", *Fraser's*, XIX, Feb. 1839, p. 163.

[28] *New Monthly*, XXI, June 1, 1827, p. 234.

V

While he was putting the finishing touches on the Serbian anthology Bowring was wondering what to do with the Polish. Finally, in despair of help from Brodziński and tired of having it around unfinished, he pieced it together himself. For introductory material he drew heavily on Szyrma's *Letters*, on Šafarik's *Geschichte der slawischen Sprache* and on Bandtkie's introduction to the dictionary supplied by Szyrma. The string of names which accounts for much of the book's Introduction Bowring got from a work on Polish prosody [30] and from Bentkowski's *Historia literatury polskiej* (1814). Bowring was under no illusion himself as to the volume's adequacy; he regretted having to send it to press in so unpolished a state, but without a working partner he was helpless. At least he refrained from dedicating it to the Tsar, as Szyrma, a sincere admirer of Alexander, urged him to do; had he done this, it is to be wondered whether the Tsar would again have rewarded him with "an amethyst ring surrounded with diamonds" [31] as he did in 1821 on the publication of the Russian *Specimens*.

Again the *London Magazine's* reviewer hailed Bowring's work. Praising his "activity" and the "universality" of his genius, he found the *Specimens of the Polish Poets* both "curious and interesting" but "by no means so pleasing to the mere lover of poetry as many of Dr. Bowring's former publications".[32] Their insipidity he attributed to Poland's "intimate connection with the rest of Europe", a connection which made them seem imitative. He liked the Lithuanian selections best, calling these the freshest and most original. Altogether this reviewer made it plain that he found the Polish verses "pleasing and graceful compositions" but "by no means either so wild or so beautiful as the specimens Mr. Bowring has given us either of the Servian or Russian".[33]

---

[29] *North American Review*, XXV, p. 355.

[30] Tadeusz Nowaczyński, *O prozodji i harmonji języka polskiego*, Warsaw, 1781.

[31] Bowring boasts of this ring on p. 123 of his *Autobiographical Memoir*, published in 1877 by his son Lewin B. Bowring.

[32] The *London Magazine* was the only important London journal to take special notice of the Polish volume. It did this in an article called "Bowring's Polish Poets", XIX, Oct. 1827, pp. 259–271. This particular quotation occurs on p. 260.

In Boston, W. B. O. Peabody, the *North American's* reviewer, found the Polish verses not "national" in the light of the strikingly "Serbian" character of the verses in the previous volume. Ten years later the Polish critic Stanisław Koźmian said virtually the same thing as the *London* reviewer, when he called the Polish verses as Bowring rendered them "somewhat spiritless".[34] It is interesting to note that the poem Peabody liked best when he reviewed the Polish *Specimens* was Brodziński's charming pastoral "Wiesław" and that this was one of the very few in the book which Koźmian felt showed "fire and inspiration" on the translator's part.

Bowring always held it against Brodziński that he had disappointed him and he never let an opportunity pass to contrast his indifference with the diligence of other collaborators. Szyrma, however, Bowring remembered with gratitude, as he did every faithful worker. He sent him a copy of the Polish *Specimens*, urged him to contribute to the *Foreign Quarterly Review*, and considered him his friend always.

VI

In the prospectus that accompanied Bowring's *Specimens of the Polish Poets* when it appeared in the middle of 1827 was "an apology for some delay in the appearance of a history of Bohemian Literature in three volumes with translations and specimens".[35] Bowring had been thinking of compiling such a work for a long time, at least since Šafarik suggested to him late in 1826 that he investigate Bohemian and Dalmatian poetry and offered to send him specimens.

Not many months later, in January of the next year, Bowring set to work in earnest on a Czech volume. His first step was to send for books and the person he chose as agent and advisor in this matter was the Slovene scholar and librarian in Vienna, Yerney Kopitar. Bowring's order the latter quickly transmitted to the venerable Czech scholar Joseph Dobrovský in Prague with whom at that time he was carrying on a vigorous scholarly correspondence. Recommending Bowring as a "famous banker" [36] and a person "about whom the newspapers often speak", Kopitar urged Dobrovský, as the only non-partisan scholar in the Empire, to assume the task of selecting the books. Dobrovský promptly sent a consignment of books and to this Kopitar himself added others, like Hanka's *Starobyla Skládánie* (1817–1824, 4 volumes), available in Vienna. Among these was a copy of Joseph Jungmann's monumental *Historie literatury české* (1825) which the author himself signed with his own hand and dedicated to "Lord Bowring", a ridiculous thing to do, Kopitar thought, but a piece of nonsense that had to stand. Late in May Kopitar finally got the package off through Bowring's Vienna agents who had been instructed to pay for the books in full.[37]

---

[33] *Op. cit.*, XXVI, Jan. 1828, p. 157.

[34] "Literature of the XIXth Century: Poland", *Athenæum*, No. 561, July 28, 1828, p. 533.

[35] *London Magazine*, XIX, p. 259.

Meanwhile, early the same spring (1827) the Bohemian folklorist František Čelakovský, having heard, presumably through Dobrovský, what Bowring intended to do, addressed himself to the latter with an offer of help. Bowring was delighted to establish partnership again with one who sounded from his letter like a good worker. His pleasure was quickly translated into action when, at very first writing, on April 3rd, he proposed to Čelakovský that they bring out a joint work. This, Bowring proposed, should consist of a history of Bohemian literature, together with two or three hundred specimens done into English verse. The specific task assigned Čelakovský was to provide material from which Bowring could form a "correct estimate of the origin, progress and present state, and general character of the popular poetry of Bohemia".

In the course of the next three months Čelakovský provided Bowring with one installment after another of Bohemian folk poetry paraphrased in German, while Bowring on his part worked diligently turning these rough fragments into English verse. "The ardent, generous and intelligent" Čelakovský, as he called him, was so satisfactory a collaborator that Bowring needed no other in a working capacity; he did yearn, however, for an important name or two which he could mention as sharing the partnership and to this end invited the much discussed Václav Hanka himself to give a few suggestions of a general character.

---

[36] *Briefwechsel zwischen Dobrovský und Kopitar*, Imperial Academy Reports, St. Petersburg, 1885, pp. 594–595.

[37] *Briefwechsel*, p. 601.

By that autumn (1827) Čelakovský, impatient for tangible results, began to urge Bowring to publish the work they had so boldly announced. Just as the verses were all in readiness except for the "final polish", publication had to be postponed indefinitely as Bowring's plans changed and he had to leave London for Heidelberg. This unexpected excursion marked the beginning of a period of readjustment in Bowring's life. The great schism of 1828 in Benthamite ranks already threatened and the complete reorganization of the *Westminster's* management was daily becoming more inevitable. Eventually Bowring induced Colonel Thomas Perronet Thompson to buy the *Review* and to hire him as editor, but months of uncertainty, to say nothing of hard feelings, intervened before the matter was settled and the attendant bitterness composed. At the same time Bowring's private affairs were nowhere nearly so prosperous as formerly; he had spent more time on a pleasant avocation than, as a business man, he could afford to spend.

As Bowring left London for Heidelberg in October, he wrote Čelakovský that he hoped to visit him personally in Prague and go over the biographic sketches for the anthology with him. But in the end the visit to Bohemia had to be abandoned and Čelakovský had to send the biographies by mail.

So far Bowring had published nothing on Bohemian literature. He now sent up, as he wrote Čelakovský, a "parachute" to precede his "balloon". The parachute took the form of an article on "Bohemian Literature" in the February, 1828, *Foreign Quarterly Review,* the journal established in the previous summer by the publishing house of Treuttel and Würtz.

"A batch of essays, each headed by the name of some foreign publication, was brought out quarterly:—this was all", lamented Robert Pearse Gillies in describing this *Review,* which turned out so differently from what he had dreamed when he conceived it. The "foreign publication" the name of which stood at the head of Bowring's article was Jungmann's *Historie literatury české,* but Jungmann's work was only a springboard for a review in the best Bowring manner. In it was little of Jungmann, a good deal of Šafarik distilled from his *Geschichte,* a large measure of opinion derived from Čelakovský and a considerable interlarding of material taken from or suggested by Anton Müller's article on Bohemian folk poetry in the August (1827) number of the *Monatschrift* published by the Bohemian Museum in Prague under the editorship of František Palacký.

Bowring's ostensible review of Jungmann seems harmless enough a hundred years later, but it got him into hot water on two counts. It involved him, in the first place, in the controversy over the manuscripts "discovered" by the man who was in 1828 librarian of the Bohemian National Museum in Prague, Václav Hanka, and it caused him in the second place to run afoul of all the Slavonic scholars in the Austrian Empire who were trying to advance the Slav cause by working in harmony with Vienna rather than in defiance of her. At the head of these stood Kopitar, splendidly ensconced in the Imperial Library in Vienna.

On September 16th, 1817, Václav Hanka had "discovered" sixteen pieces of what seemed to be XIIIth century Bohemian poetry, eight verses of which were epic in spirit, the rest love songs. Calling these, from the place of their discovery, the *Queen's Court Manuscript* (*Královédvorský Rukopis*), Hanka published them in 1818 in the original old Czech and the next year brought out a German version of them with the help of his friend Václav Alois Svoboda.

Some time in 1819 two more Bohemian poems alleged to be of still more ancient vintage than the *Queen's Court* verses were sent to Burgrave Franz Kolowrat for the Bohemian Museum. These two, "The Assemblies" (Sněmy) and "Libuša's Verdict " (Libušin Soud), called collectively the *Green Mountain Manuscript* (*Zelenohorský Rukopis*), were published a year or so later in the new Warsaw journal *Prawda Ruska.* In 1822 they were reprinted by the brothers Jungmann in their new periodical *Krok.*

When the *Queen's Court Manuscript* appeared it caused a stir in Slavonic circles and there was more or less general agreement that here was indeed a splendid find. When the second group of verses was published divided opinion greeted it. Joseph Dobrovský, the most respected Slavonic scholar in the Empire, pronounced them the ingenious work of an impostor and in this opinion he was heartily seconded by Kopitar. Opposed to this school was a group that held the manuscripts to be genuine, with the very Václav Hanka whom Kopitar accused of having been the actual "scripturae felix imitator",[38] at its head. In this camp stood not only Bowring's collaborator and friend Čelakovský but Jungmann and Šafarik and Palacký as well.

Matters stood thus when Bowring rode onto the field. He would have done better, Dobrovský hinted, to confine himself to money matters than to enter the philological lists.[39] Not Bowring. He did in his article for the *Foreign Quarterly Review* exactly what he told Čelakovský he would do; he gave both sides. He had early been advised by Šafarik that the "best of Bohemian verse" was to be found in Hanka's collection and in this opinion had been reinforced by Čelakovský. Dobrovský, on the other hand, had called them the work of an impostor. Taking a stand of his own on middle ground, Bowring drew attention to the fact that Dobrovský had indeed called the verses spurious but that "a contrary doctrine" was, however, "held by Hanka, Čelakovský, and other authorities".[40]

It was high summer and Dobrovský was at Chudenitz when he saw Bowring's article. So this was what the Englishman had made out of that package of books they had sent him, he wrote Kopitar. He had not only dared throw suspicion on the scholarly judgment of himself, Joseph Dobrovský, but, worse than that, had set up Hanka and Čelakovský as "authorities" of the same stature as himself![41]

Dismissing the review as Bowring's "Dandy-Artikel",[42] Kopitar charged into battle. He was hurt and angered for his beloved master's sake and he lost no time in telling Bowring exactly the impression his article had made. "You have insulted the Abbot Dobrovský", he wrote, "the Abbot who is a venerable sage, the greatest *savant* and critic of literature and history not only of the Slavs but of all Austria, and against him you cite two 'blanc-becs' like Č and H, pupils of his, men entirely without authority among scholars in general."

Dobrovský first, then Kopitar in his turn, were both much more than angered by Bowring's article; they were seriously disturbed by its implications. Bowring's repeated allusions to the anti-Slav activities of the Vienna police and of the Imperial Government itself made both uneasy. Dobrovský saw at once the danger in Bowring's "invectives against the Austrian government" [43] as he called them, while Kopitar charged Bowring straight out with perpetrating untruths. He denounced him for painting "our police like the Spanish Inquisition" by declaring that "The very names of Bohemia and Bohemians it is the policy of the government of Bohemia to banish" and other similar statements.[44] To Bowring's observation, provoked by Kollár's "Daughter of Slava", that "We have been somewhat surprised that language so free and sentiments so lofty should have been allowed to circulate in Bohemia" [45] Kopitar protested that Kollár's enemies would certainly use Bowring's words to prove Kollár's designs criminal, especially Bowring's sly comment that the "dangerous nature" of Kollár's thoughts must have been hidden from the censor by the poem's allegorical trappings. All Bowring could do to repair the damage he had done, Kopitar warned, was to refrain, at least, from repeating such words in the final Anthology.

---

[38] *Briefwechsel*, p. 597.
[39] *Briefwechsel*, p. 617.
[40] *Foreign Quarterly Review*, II, p. 149.
[41] *Briefwechsel*, pp. 616–617.
[42] *Briefwechsel*, p. 619.

Čelakovský was no less anxious over the article's possible effect on Kollár's fortunes than Kopitar. He liked the review on the whole but this aspect of it worried him; how would Vienna take it? To his worried query about this regrettable consequence of a well-intentioned article, Bowring replied that he was confident no harm would come of it, since all who read the article would undoubtedly be lenient in their judgment of it. "We say many things without weighing them," he wrote Čelakovský, "because nobody calls us to account for what we say. But we must not be judged harshly, for we mean well. We desire, I hope, so to see you all as happy and as free as may be, at least as free as would add to your happiness, and we have to tune our harps to English feelings and take a somewhat advanced stand in the political contest."

Kopitar later toned down his rebukes to Bowring but he never departed from the position he took originally that Bowring should realize the harm he could do the Slavonic cause "by heaping charges of injustice on Vienna". In the manuscript controversy it was Bowring who did not depart from the position he had first taken; his final word in the *Cheskian Anthology* was, "Between such authorities I dare not attempt to decide." [46]

In 1829, when Bowring would normally have published the

---

[43] *Briefwechsel*, p. 616.

[44] *Foreign Quarterly Review*, II, p. 147.

[45] *Foreign Quarterly Review*, II, p. 167.

[46] *Cheskian Anthology*, p. 9.

Anthology in the projected three-volume form, the firm in which he was a partner failed, and he was obliged to seek a post in the government service. He had, therefore, to keep putting Čelakovský off from month to month, finally from year to year, urging him at length that if he wished to see the Anthology published he would have to secure at least a hundred subscribers in advance, since he himself, the father of seven children, was in no position to risk a loss and since the temper of the times was now one of indifference to everything except politics.

To keep alive popular interest in Bohemian poetry while the book was hanging fire, Bowring sent up a second parachute with his review of Hanka's spectacular discoveries in "Ancient Bohemian Ballads", published in the *Westminster Review* for April, 1830. In the late twenties Bowring's work had taken him to Holland and in 1829 he had been awarded an honorary doctorate in languages by the University of Gröningen. Subsequently he was sent on a mission to Paris and his research in that city kept him for months extremely busy. In the middle of 1831, however, when Čelakovský finally came across with forty subscribers to the Anthology, sixty less than stipulated, Bowring took time to put the book through the press on the strength of this guarantee. Thus, in a manner totally different from that foreseen in 1827 when it was undertaken, the fourth and final Slavic Anthology saw the light of day. It appeared in the year of the great Reform Bill, a modest volume just a third as large as first proposed, with a warm dedication to the one who had made it possible, František Čelakovský.

The only important periodical to review the *Cheskian Anthology* specifically was the weekly *Athenæum*, a journal of strong Liberal tendencies, friendly on the whole to Bowring. By this time Bowring had published translations from Russian, Batavian, Spanish, Serbian, Polish, Dutch and Magyar, tongues so numerous and unfamiliar, not to say amusing, to the average English reader that a new volume, and that a "Cheskian", was an event not unlikely to provoke a wondering "What next?". In this vein wrote the *Athenæum* on the 31st of March:

> We lie under personal obligation to Mr. Bowring—we beg his pardon—he is a Doctor, if not of Laws, at least of Languages. We knew him—or rather he knew us—in infancy, when he had the kindness to translate our little wants from the Baby-lonian into the mother-tongue. In our school days he volunteered to do our exercises in French, Latin and Italian, and was our proxy, we remember, in learning Greek and Hebrew. . . .

> These are private obligations, but Mr. Bowring has added to our national debt to him by publication of the *Cheskian Anthology*. The poets of Britain must rejoice to find that they have such a band of Bohemian Brothers as sing in this little volume. . . . Poems done into English generally drink dreadfully like the homemade wines—they may be named after the Spanish or the Rhenish—but they smack of nothing but domestic currant and gooseberry. This is not the case with Mr. Bowring. He imports or smuggles over the genuine spirit of his author—Spanish or Polish, Russian or Magyar . . . etc.
>
> We would fain quote a few of the early lyrics and some beautiful sonnets from Kollar, but want of room forbids, and besides, we can safely advise the lover of poetry to extract the whole volume from Hunter.[47]

VI

The *Athenæum* was by no means the only, nor even the first, London journal to find Bowring's fecundity and ubiquity as a translator amusing. It was *Fraser's Magazine* that started the vogue of poking fun at the good Doctor and even made it "the thing" to do so. As a matter of fact it would not have taken the eight assorted volumes of poetry that had appeared by 1830 to set *Fraser's* off: if this tumultous journal had existed when the very first of them, the Russian, appeared in 1821, Bowring would undoubtedly have been summoned then and there to one of Nol Yorke's fictitious levees for de-frocking. As it was, the Fraserians waited only until their second meeting, in March 1830, to loose their barbs. Teasing a Benthamite was such good fun that Maginn and his fellow-wits were loath to let their victim go until they had given him a good trouncing and stripped him of what they called his garb of pretense. To prolong the sport and to give their jeers a pontifical air not unlike Bowring's own, the Fraserians fabricated a reviewer, a polyglot more accomplished even than Bowring, to deal with the Doctor's verses. This fellow, whom they called by the absurd name of Quaffypunchovicz, they turned loose on poor Bowring![48]

Less in Fraserian raillery than in pure fun was the *Metropolitan's* critique of Dr. Bowring in the Fable of the Dog that Talked.[49]

---

[47] *Op. cit.*, No. 231, p. 203.

[48] *Fraser's*, I, Mar. 1830, pp. 155–177; May, 1830, pp. 433–442; June 1830, pp. 601–604.

[49] "Discovery of a New Language by Dr. Bowring", by Mark Marvel, *Metropolitan Magazine*, VI, Feb. 1833, pp. 148–153. The Dog-Rib Indians were in the news at this time, thanks to the publication in 1828 of the *Second Narrative* of the Franklin Polar Expedition which mentions them extensively.

"On a drizzly, damp November night" the author of the fable was "startled out of his preoccupation" as he was returning home through a heavy London fog by the pitiful moan of a dog. Shocked by the poor creature's pinched look, he did what he could to relieve its hunger by feeding it bits of gingerbread that he found in his pocket. Suddenly, in a moment of the dog's "suspended mastication", the author detected something that startled him: the noises the poor thing emitted "wore the semblance of articulate, syllabic sound"! Here Dr. Bowring entered—for he alone of all the author's acquaintance was competent to communicate with this "singular quadruped". Fortunately Dr. Bowring was in the neighborhood. Tying the dog to a post, the author routed out the learned Doctor and led him back to the Dog. In no time at all Dog and Doctor were engaged in animated conversation! Dr. Bowring had *placed* the dialect, obscure though it was, spoken by the unhappy creature: it was a "modification of the dialect spoken by the *Dog-ribbed Indians* of North America"! The upshot of the encounter was, of course, that Dr. Bowring undertook "a volume on the language and poetry of dogs and its affinity with that of the *Dog-ribbed Indians*"!

The most friendly gibe at Bowring's polyglottism was a doggerel by his colleague in the Liberal movement, Thomas Hood:

To Bowring! man of many tongues,
   (All over tongues, like rumour)
This tributory verse belongs,
   To paint his learned humour,
All kinds of gab he knows, I wis,
   From Latin down to Scottish—
As fluent as a parrot is,
   But far more Polly-glottish.

No grammar too abstruse he meets,
   However dark and verby;
He gossips Greek about the streets,
   And sometimes Russ—*in urbe.*
Strange tongues—whate'er you do them call,
   In short the man is able
To tell you what's o'clock in all
   The dialects of Babel.

Take him on 'Change—in Portuguese,
   The Moorish and the Spanish,
Polish, Hungarian, Tyrolese,
   The Swedish or the Danish;

Try him with these and fifty such,
His skill will ne'er diminish,
Although you should begin in Dutch,
And end (like me) in Finnish.[50]

Bowring received, however, his share of serious attention and honest appraisal even after he had published so numerous a string of anthologies as to make even his most partisan supporter suspect him of what *Fraser's* called "nineteenth-centuryizing us, with this same knack of translating from all the tongues of Babel".[51] *Fraser's* itself, for all its raillery, had to admit that "In the province of song and ballad translation probably few men have deserved better than Dr. Bowring; not on account of any supreme manifestation of talent in the individual, but simply on the score of industry." [52] In 1831 the *Edinburgh Review*, having weighed the merits of all Bowring's translations as a whole, came to this conclusion about the many-tongued Doctor:

> His varied and almost Mithridatic acquaintance with the languages of modern Europe, extending even to their less classical or almost forgotten dialects, and that liberal spirit in literature, which so extensive a field of enquiry is sure to produce, seemed peculiarly to mark him out as one fitted to transfer to his country those strains which had conferred celebrity on their authors in *their* own; or which, though their origin and authorship are lost in the darkness of antiquity, had long cheered the peasant in his sledge amidst the frozen snow, or been associated with the jollity of the harvest and the vintage, or the more tranquil mirth of the fireside.[53]

The two contemporaries most competent to pass on Bowring's merit as a translator and on the value of his contribution in the field of Slavonics were both across the Atlantic in New England. The two were Talvj, now Mrs. Edward Robinson, and the poet James Gates Percival, the former a resident of Boston, the latter one of New Haven's "characters". Talvj appreciated Bowring's pioneering; his translations and surveys were in every case the only ones in English to which she could refer her readers when she herself undertook in America a pioneer venture not unlike that Bowring had begun in England more than a decade earlier. In the editorial work on which he was engaged in the late 1820's Percival drew heavily on the Introductions to Bowring's volumes and in the early 1830's, moreover, when he began to translate from the Slavonic tongues himself, he followed Bowring's guidance almost slavishly. Percival did, however, what Bowring never did; he freed himself quickly from the German intermediary tongue and mastered Russian and Serbian, Polish and Czech sufficiently for translation from these languages directly. In the end, therefore, Percival turned on his first guide and rebuked him for his superficiality.[54]

---

[50] Lewin B. Bowring, *Autobiographical Recollections*, London, 1877, pp. 63–64.
[51] *Fraser's*, I, May 1830, p. 433.
[52] *Idem*, Mar. 1830, p. 158.
[53] *Op. cit.*, Jan. 1831, pp. 322–323.

Another contemporary who, being something of a linguist himself, had a good deal to say about Bowring as a translator was the American John Neal. In the *Knickerbocker Magazine* (New York) for November, 1833, Neal wrote,[55]

> As a translator we think rather highly of Dr. Bowring. His method is peculiar and the results happy, though we are inclined to believe his translations of great authors mere paraphrases and unfaithful in what should be the predominating features of such labors. The music is preserved and the thoughts, when those thoughts are on a level with the translator's imagination. But beyond this they are of necessity feeble and false; and as we can easily measure the height and depth, yea, the length and the breadth of Dr. John Bowring's imagination by his original pieces, we may be sure that the burning passion, the overpowering pathos and the terrible sublimity of the great northern bards *are not* and *cannot be* rendered by him.

Neal did not like Bowring at all, yet he was obliged to admit Bowring's enormous reputation, "not only at home, but here [in America] and throughout the whole of northern and a part of southern Europe; now as a critic and reviewer and now as the author of much beautiful and simple poetry; here as a linguist hardly inferior to Sir William Jones himself and there as the only 'faithful and free' translator of many an unknown language, with all its treasures and glories. . . ."[56]

Such a reputation Bowring had, undeniably, both on the continent and at home, as well as in America. Few of Bowring's admirers realized, even William J. Linton who called his translations "too facile",[57] that everything he versified in English had come into being, as Shaw pointed out, by being distilled through the "alembic of an intermediary tongue".[58] Nor would most of Bowring's readers—Carlyle and Thackeray and the other Fraserians excepted—have found any fault with the method, had they known it. Bowring's generation did not take the literature of the Slavs very seriously. Even German was only beginning to be treated as a subject for earnest study, leading the procession in which later the Slavonic tongues were to march. Bowring was the sort of interpreter of the Slavs his generation could understand. His very grandiosity was a recommendation, his *Schwärmerei* attractive.

---

[54] See letters and unpublished writings of James Gates Percival, Yale University Library, New Haven, Conn.

[55] "Portraits of Distinguished Contemporaries: Dr. Bowring", *op. cit.*, II, Nov. 1833, p. 365. Neal wrote this in a pseudo-Fraserian manner.

[56] *Idem*, p. 358.

[57] William James Linton, *Memories*, London, 1895, p. 160.

The final judgment of Bowring's contemporaries upon him as a translator of Slavonic verse is well summarized in the *Edinburgh Review's* opinion, delivered at the time of reviewing all his Slavonic volumes except the Cheskian:

> The translator is to poetry what the adventurous merchant is to commerce. He circulates the produce of thought, varies our intellectual banquets, teaches us that some accession to our store may be derived from those quarters which we had regarded as the most sterile and unpromising, and thus adds another link to the chain of social and kindly feeling which should bind man to his fellows. In this commerce of mind few have labored more assiduously than Dr. Bowring.[59]

Linton said there was about him "a little of the Girondist, of the pedagogue", but that "he was a good citizen and a man to be respected".[60] He was exactly that and an enterprising indefatigable peddler besides.

Chronological List of Writings on Slavonic Poetry by John Bowring himself in British Periodicals during the years 1821–45

1. "Politics and Literature of Russia", *Westminster Review*, I, pp. 80–101, Jan. 1824.
2. "Specimen of Popular Poetry from the Old-Sclavonic-Polish Dialect spoken in the Province of Volhynia" [*The Three Fountains*], *London Magazine*, IX, pp. 132–133, Feb. 1824.
3. "From the Polish of Zimorowicz" [*I saw thee from my casement high*], *London Magazine*, IX, p. 188, Feb. 1824.
4. "J. A. Krilov's Russian Fables", *Westminster Review*, IV, pp. 176–178, July 1825.
5. "Popular Poetry of the Servians", *Westminster Review*, VI, pp. 23–39, July 1826.
6. "Bohemian Literature", *Foreign Quarterly Review*, II, pp. 145–174, Feb. 1828.
7. "Illyrian Poems", *Westminster Review*, X, pp. 71–81, Jan. 1829 [*Probably* by Bowring].
8. "The Present State of Literature in Poland", *Foreign Quarterly Review*, V, pp. 699–704, Feb. 1830.
9. "Ancient Bohemian Ballads", *Westminster Review*, XII, pp. 304–321, April 1830.

---

[58] See article cited in Note 2, p. 282.

[59] *Op. cit.*, Jan. 1831, p. 322.

[60] Linton, *Op. cit.*, p. 160.

Chronological List of Articles containing critical comment on John Bowring as a translator of Slavonic Poetry, 1821–1845

1. "Russian Poetry", *London Magazine*, III, pp. 316–321, March 1821. [A review of Vol. I of *Specimens of the Russian Poets*]
2. "Servian Popular Poetry", *London Magazine*, XVII, pp. 567–583, April 1827. [A review of Bowring's *Servian Popular Poetry*]
3. "Servian Popular Poetry", *New Monthly Magazine*, XXI, pp. 232–234, June 1827. [A review of Bowring's *Servian Popular Poetry*]
4. "Bowring's Polish Poets", *London Magazine*, XIX, pp. 259–271, Oct. 1827. [A review of Bowring's *Specimens of the Polish Poets*]
5. "Servian Popular Poetry", *North American Review*, XXV, pp. 352–366, Oct. 1827. [A review, by Jared Sparks, of Bowring's Serbian volume]
6. "Russian Literature", *Foreign Quarterly Review*, I, pp. 595–629, Nov. 1827. [A review, by J. G. Cochrane and Victor Smirnov, of Bowring's translation of Zhukovsky's "Svetlana"].
7. "Specimens of the Polish Poets", *North American Review*, XXVI, pp. 146–157, Jan. 1828. [A review, by W. B. O. Peabody, of Bowring's Polish volume]
8. Indirect References in
   "Poetry of the Magyars", *Fraser's Magazine*, I, pp. 155–177, Mar. 1830.
   "The Magyars vs. Dr. Bowring", *Fraser's*, I, pp. 433–442, May 1830.
   "Kisfaludy's Meeting of the Similes", *Fraser's*, I, pp. 601–604, June 1830.
9. "Dr. Bowring's Poetical Translations", *Edinburgh Review*, LII, pp. 322–337, Jan. 1831. [A review of all these except the Polish volume]
10. Review by W. H. Leeds of Batiushkov's *Opyte v' Stikhakh i Prosye*, 2 Vols., St. P., 1830, *Foreign Quarterly Review*, IX, pp. 219–222. [Contains comment on Bowring's Russian volume]
11. Review of Bowring's *Cheskian Anthology*, *Athenæum*, No. 231, p. 203, Mar. 31, 1832.
12. "Discovery of a New Language by Dr. Bowring", *Metropolitan Magazine*, VI, pp. 148–153, Feb. 1833. By Mark Marvel.
13. "Portraits of Distinguished Contemporaries: I. Dr. Bowring", *Knickerbocker Magazine*, New York, II, pp. 358–368, Nov. 1833. By John Neal.
14. "Two Articles on the Annuals", *Fraser's*, X, pp. 614–615, Nov. 1834. By John Churchill.
15. "The Poems of Quaffypunchovicz", *Fraser's*, XIII, pp. 593–599, May 1836. By John Churchill.
16. "Literature of the XIXth Century: Poland", *Athenæum*, No. 561, July 28, 1838, pp. 532–536. By Stanislas Koźmian. References to Bowring, p. 533.
17. "Fabulists of Russia", *Fraser's*, XIX, pp. 153–163, Feb. 1839. Reference to Bowring in footnote to pp. 159–160.
18. "The Past and Present of Servia", *Foreign Quarterly Review*, XXIX, pp. 1–31, April 1842. [A review of Ranke's *Die serbische Revolution* and Bowring's *Servian Popular Poetry*]
19. Introduction by Thomas B. Shaw to his translation of *Amalát Bek*, *Blackwood's Magazine*, LII, pp. 282–283, March 1843.

SYMPOSIUM
ON
RECENT ADVANCES IN PSYCHOLOGY

Papers read before
The American Philosophical Society
Annual General Meeting
April 25, 1941

PHILADELPHIA
THE AMERICAN PHILOSOPHICAL SOCIETY
Independence Square
1941

SYMPOSIUM
ON
RECENT ADVANCES IN PSYCHOLOGY

Coalescence of Neurology and Psychology 461
KARL S. LASHLEY, Research Professor of Neuropsychology, Harvard University.

The Genesis of Behavior Form in Fetus and Infant 471
ARNOLD GESELL, Director of the Clinic of Child Development, Yale University.

On the Nature of Associations 489
WOLFGANG KÖHLER, Professor of Psychology, Swarthmore College.

Mental Abilities 503
EDWARD L. THORNDIKE, Professor of Educational Psychology, Teachers College, Columbia University.

Psychoanalysis and Scientific Method 515
CARNEY LANDIS, Principal Research Psychologist, New York State Psychiatric Institute; Associate Professor of Psychology, Columbia University.

Psychology and Defense 527
ROBERT M. YERKES, Professor of Psychobiology, Yale University.

Motivation, Learning and Adjustment 543
EDWARD C. TOLMAN, Professor of Psychology, University of California.

COALESCENCE OF NEUROLOGY AND PSYCHOLOGY

KARL S. LASHLEY

Research Professor of Neuropsychology, Harvard University

(Read April 25, 1941, in Symposium on Recent Advances in Psychology)

Studies of the nervous system and of physiological psychology have as a common aim the understanding of mental processes in terms of the activities of the brain. In his Gifford lectures, delivered in 1938, Sir Charles Sherrington presents an extensive analysis of the relations between events within the brain and the phenomena of mind (12). Summing up the problem he says:

> No attributes of 'energy' seem findable in the processes of mind. That absence hampers explanation of the tie between cerebral and mental. Where the brain correlates with mind, no microscopical, no physical, no chemical means detect any radical differences between it and other nerve which does not correlate with mind. In both regions, whether 'mental brain' and [or] 'non-mental brain', changing electrical potentials along with thermal and chemical actions make a physiological entity held together by energy-relations. To correlate with that physiological entity, a suite of mental experience, a complex of thought, feeling, conation, an activity no doubt, but with what if any relation to electrical potential, heat and chemistry. For myself, what little I know of the how of the one does not, speaking personally, even begin to help me toward the how of the other. The two for all I can do remain refractorily apart. They seem to me disparate; not mutually convertible; untranslatable the one into the other (page 312).

With this statement the greatest living neuro-physiologist despairs of finding a common ground between the sciences of the brain and of mind. I am confident, however, that the outlook is not so hopeless as Sir Charles believes. In fact he seems to have missed a solution of the problem by no more than the turning of a page. In the same lectures he has faced the problem of the nature of life and has found that life is not a thing attached to this or that substance or chemical action, but is organized activity, varying in character with complexity of structure and ranging without discoverable discontinuity from the nearly crystalline simplicity of the filterable virus to the elaborate organization of the mammalian body. He has just missed seeing that mind also is not a thing attached to life, a unique form of existence, but is a term including an indefinite number of complex structures or relations.

Perhaps the most important contribution of psychologists to this problem has been the realization that the characteristics of the mental can be stated meaningfully only as a structure or organization of elements which are themselves as purely conceptual as is the energy of physics. Such a notion was foreshadowed by the growth of behaviorism (5, 7), but it remained for the logical positivists to develop a critique of scientific thinking which gives it rigorous formulation (13). When the supposed characteristics of the mental are tried in the fire of operational definition, the most imposing of them evaporate. Mind, when analyzed to its definable constituents, has no discernible properties other than organization or integration of processes which differ through a range of complexities as wide as the structural differences between the virus and the human body.

The task of understanding how the brain thinks is simplified by such an analysis of mind and there are already indications that the electrical and chemical activities of the brain are organized in just those ways which have been thought to characterize as unique the organization of mind. As a leading student of electrophysiology has recently said, "The behavior of brain-waves seems more psychological than physiological". There is no need to seek, as Sherrington has done, a fundamental difference between "mental" and "non-mental" brain, to look for a special form of energy or of chemical action correlating with mind, but only to look for correspondences of organization in physiological and mental processes.

In the brief time here I can do no more than indicate some of the areas of investigation where physiological and psychological studies are revealing similar complexities among physical and mental events. There is not yet a complete coalescence of the neurological and psychological phenomena at any point, but there are regions of confluence where the transition from a metaphysical to a scientific, experimental approach to the problem of mind and body is clearly foreshadowed.

Studies of the nervous system have advanced most rapidly in the analysis of the transmission of nerve impulses and the integration of spinal reflexes. Generalization from such studies has produced a picture of the nervous system as a great assemblage of reflex connections in which nervous impulses are transmitted over limited paths in one direction only from sense organ to muscle. It is this picture which is difficult to harmonize with the facts of psychology.

Recent anatomic and physiologic studies however give a quite different conception of the integrative processes within the central gray matter. An advance of prime importance was the report by Lorente de Nò (10) of the intricacy of connections between the cells of the hippocampal cortex. Instead of one-way transmission over limited paths, his figures show the anatomic possibility of diffuse spread of excitation through nervous tissue, almost as through a continuous network. In addition, return circuits make possible a recurrent excitation which will maintain activity for various lengths of time and, of course, greatly complicate the pattern of excitation. The action of such recurrent or reverberatory circuits has been demonstrated experimentally (11).

The physical properties of such a system are quite different from those of restricted reflex circuits. If the spatial pattern of incoming impulses is large in comparison with the units of nervous organization, as is the case, then the cortex will have some of the properties of a homogeneous medium for transmission of the effects of stimulation. Excitation at a point will spread as a wave of nervous impulses traversing the network in all directions. Such waves will reinforce or cancel each other according to their time relations and the refractory periods of the transmitting system. With repeated volleys of impulses at regular intervals the activity of the cortex will tend to stabilize in a repetitive pattern extending throughout the sensory area. The analogy with wave transmission on the surface of a liquid is fairly close, although any excitation pattern must be greatly modified by the inherent characteristics of the reverberatory circuits.

Studies of the electrophysiology of the cortex lend some support to this inference from structure. In a quiescent state the nerve cells of large cortical areas fall into synchronous activity which is detected through their summated action potentials, recorded in the so-called brain-waves of the electroencephalogram. This activity appears as successive fluctuations of electrical potential at frequencies of from 4 to 30 per second, the frequencies relatively constant for each part of the cortex and maintaining a uniform character over entire cortical areas (2).

One of the chief stumbling blocks for neurological theory has been the fact that whereas the reflex seems determined by certain definite nerve cells, psychological functions are largely independent of such definite structural limitations. For example, an object seen with one part of the retina is immediately recognized when its image falls on another part, activating an entirely different group of nerve cells. It may, indeed, be stated as a general principle of behavior that psychological functions are of such a character that they must be independent of the particular nerve cells excited and be determined by the patterns or relations between points of excitation (8).

Much of my own work, based on operative destruction of parts of the brain combined with experimental studies of learning, has shown a similar independence of function from the structural details of the brain. Thus a differential reaction to visual stimuli will survive the operative destruction of any part of the brain except the lateral margin of the visual cortex and, in addition, any part of this critical area may be destroyed and the visual organization will still persist. The individual elements of the visual cortex are equivalent for the visual reactions. Similar results for other functions and other parts of the brain have been obtained (9).

Such an organization can be understood only in terms of a reduplication of functional elements. The equivalence of parts which is apparent in the growth of crystals or of the early blastula involves a reduplication of molecular elements or of chromosome mechanisms. The structure of the cerebral cortex seems to provide in the same way for a reduplication of functional elements. Although little is as yet known of the details of nervous integration within the cortex, such a system as I have described implies an interference pattern of propagated impulses which may be reduplicated throughout an entire functional area and so induce functional effects which are independent of the particular structural elements transmitting the pattern. Such a model meets the requirements imposed by psychological data and from it also some of the facts of perceptual organization can be predicted.

Students of the psychology of perception are approaching this same problem by different methods. Analysis of the ways in which the sensory elements are combined in perceptual patterns reveals that the laws of their combination correspond closely to the physical laws of energy distribution. For example, the mutual influences between a perceived figure and the background on which it occurs are in part comprehensible in terms of interphase phenomena at the boundaries between the figure and its background. The degree of interaction of separate portions of the visual field varies with their distance of separation in ways characteristic of a gradient system (6).

The development of this subject, which is due largely to Professor Köhler and his associates, represents an important step toward a reduction of perceptual organization to physiological terms. It is true that these two lines of analysis, neurological and psychological, have not yet led to a common description of organization. The behavior of perceptual elements is for the most part best described in terms of the interaction of forces in a homogeneous field, in terms of chemical gradients, of potential differences or the like, whereas the patterns of nervous activity reveal themselves rather as interference phenomena among propagated waves of excitation. The important point at the present stage of the science, however, is that the analyses of organization show a convergence in which the electrochemical activities of the brain are seen to have some of the properties of mental organization, and the behavior of percepts, some of the characteristics of a physical system. Such progress is ground for the hope that a common statement of principles may be reached in the not too remote future.

A second point of convergence of physiology with psychology is in the problem of motivation. Psychology speaks of emotional tensions, drives, interests, and the like, and is compelled to express these as quantitative variables representing forces of unknown character, which are not directly discoverable in experience. Explanation of them is sought in physiological processes, either explicitly, with the recognition that they have no mental characteristics, or implicitly by the postulation of an unconscious mind, as in the psychoanalytic fantasies. Physiological studies are making some progress toward an understanding of the nature of these forces, especially in the fields of sexual behavior. Three variables have been found to contribute to the dominance of a pattern of behavior. These are chemical activation, adequacy of the stimulus situation, and amount of available nervous tissue.

The sexual or maternal drives can be initiated or strengthened by the injection of specific hormones. The sexual interest and activity of the castrated male animal is restored by injection of testosterone. Waning maternal interest may be strengthened by an extract from the hypophysis. This chemical reinforcement of the drives is specific for several drives and hormones, so it is most probable that the substances act directly upon the nervous system to increase the excitability of distinct nervous organizations, which however do not have a restricted localization.

Destruction of parts of the cerebral cortex reduces the intensity of the sexual drives, somewhat in proportion to the extent of damage. Males with more than half of the cortex removed may show no reaction to a receptive female. Large doses of testosterone, however, will restore them temporarily to normal interest and activity. There is some evidence that the dose of hormone required is directly proportional to the amount of brain tissue destroyed (1).

Finally, the excitability and intensity of the drive varies with the adequacy of the stimulating situation. A nursing mother rat weans her young and ceases to care for them when they are about 18 days of age. If at that time she is given a new-born rat, she not only takes it to her nest and nurses it, but she also shows a revived interest in the youngsters that she has been weaning and drags them also back to her nest. The same revival of the maternal drive has been achieved by hormone injections (15).

Thus in these primitive drives chemical, environmental, and central nervous factors are somewhat interchangeable. Weakening of any one of the three may be compensated for by addition to the others. The ultimate mechanism is the nervous organization of which the excitability may be increased by the hormone, by appropriate stimulation, or by some factor which correlates with the amount of nervous tissue intact. This latter factor is most probably some mutual reinforcement of equivalent patterns of nervous activity. There is evidence from other sources that, when a system the parts of which are equivalent for a reaction, as was described for the visual cortex in discrimination of visual patterns, is destroyed, the efficiency of what remains is reduced, although its functional organization is retained. So apparently the intensity of a drive may be diminished by a brain injury, without abolishing its organized activity.

Physiological studies have as yet made no progress toward analysis of the less primitive motives of human behavior. In human conduct social influences, that is, effects of training, are far more dominant than in animal behavior. Such training builds up enormous complexes of interrelated habits which have some mutual facilitating effect. The dynamic mechanism so established is probably comparable to the central nervous factor in the more primitive drive, where the mass of excitable tissue contributes to the intensity of the drive. Professor Woodworth (16) has marshalled psychological evidence in support of the interpretation of human motivation as an accumulation of such interrelated habit systems, exerting a mutual reinforcement and thus acquiring the potency of a drive. The neurological evidence is consistent with such a view.

There is thus a general recognition that the dynamics of behavior is not available for direct introspective study, that it is nonconscious, or non-mental, and can only be reached by inference from psychological studies. Meanwhile, physiology is making some significant progress toward an understanding of the actual processes involved in motivation at least at a primitive level.

A third point of convergence of neural and psychological analysis is on the problem of the organization of intellectual functions. Psychological analysis of the components of intellect has been largely a logical analysis of its achievements. That is, intellectual functions have been classified not in terms of how they work or of how they are related, but in terms of what they accomplish. The mental faculties of classical psychology were of this character and even today such terms as emotion, perception, imagination, abstraction, or reasoning do not represent functional groupings of processes of like nature, but only classes of such activities as achieve comparable results.

It has never been possible to understand or to classify the mental defects resulting from brain injuries in terms of such psychological categories. The patient with a temporal lesion shows limited disturbances in speech and in comprehension of some logical relations which have not corresponded to any recognized unitary psychological functions. The primary defect may be in the formation of individual words, as in Head's verbal aphasia (4), or in memory for names of common objects, as in amnesic aphasia (3), or in understanding of the spatial relations when stated in verbal form (4). Such defects of the language function are unintelligible in terms of the current psychology of language, and point to physiological variables which do not correspond to current classifications.

In the past decade a new tool for the analysis of intellectual functions has been developed, chiefly by Spearman in England and Thurstone in this country. A large number of people are put through a great variety of performance tests, in fact, the attempt is made to test their abilities in practically everything that a man can do. The scores are then subjected to a vector analysis to determine from the clustering of variations in what tests individuals tend to be uniformly good and in what uniformly bad. The intercorrelation of the test scores indicates those tasks which are dependent upon a common ability or a specific kind of performance, and so lead to a definition of independently variable mental functions.

This method is discovering functional variables which are not those of classical psychology. For example, it distinguishes between the ability to think in terms of the spatial relations of objects and to comprehend non-spatial relations. It distinguishes facility in manipulating separate symbols, as in recognition of words in jumbled letters, from facility with combinations of symbols as in grammatical speech (14). It is difficult to describe these variables except in terms of the tests from which they are derived, for they do not correspond to any familiar classification of functions. They do seem, however, to correspond to functions which may be independently lost as a result of localized brain injury. Certain types of apraxia are marked by difficulty in dealing with spatial relations; the function represented by manipulation of isolated symbols resembles the ability which suffers in verbal aphasia as defined by Henry Head, and there are other less clear correspondences.

Psychology has still to discover how the various factors revealed by such analysis interplay to produce organized thought. Neurology likewise still has much to do in the investigation of the interaction of cortical fields which are associated with diverse functions. Nevertheless the discovery that the various capacities which independently contribute to intellectual performance do correspond to the spatial distribution of cerebral mechanisms represents a step toward the recognition of similar organization in neurological and mental events.

I have chosen to deal here with progress toward a remote goal, which must be treated in somewhat general terms, rather than to report more conclusive evidence on a limited problem, because the ancient question of the relation of mind and body underlies all neurological and psychological investigations and because the developments of the last decade seem to me to point to a possible solution, as earlier work on nerve physiology and cerebral localization did not.

The recognition that the mental has no characteristic properties other than organization, which is arrived at by critical examination of the concepts of psychology, paves the way for a comparison of neurological and mental organization. Studies of the anatomy and physiology of the cerebral cortex are revealing integrative mechanisms which differ significantly from the reflexes regarded as the elements of nervous organization by physiologists a decade ago. The physiological organization has some at least of the characteristics of mental organization. Analysis of mental phenomena is likewise converging toward a description of organization identical with the physiological. Just as, with the recognition of the complexity of the atom, physics has progressed to a reduction of the apparent qualitative diversities of the chemical elements to differences in atomic structure, so neurology, by discovering the complexities of cerebral organization, seems to be opening the way for reduction of the qualitative diversities of mental states to terms of the physiological activities of the brain.

BIBLIOGRAPHY

1. BEACH, F. A. 1940. Effects of cortical lesions upon the copulatory behavior of male rats. *Jour. Comp. Psychol.*, 29, 193–245.
2. GIBBS, F. A., AND GIBBS, E. L. 1941. *Atlas of Electroencephalography.* Cambridge. Pp. 1–221.

3. Goldstein, K. 1937. The problem of the meaning of words based upon observation of aphasic patients. *Jour. Psychol.*, **2**, 301–316.
4. Head, Henry. 1926. *Aphasia and Kindred Disorders of Speech.* (2 vols.) London. Pp. xvi + 550; xxxiv + 430.
5. Holt, E. B. 1914. *The Concept of Consciousness.* New York. Pp. xvi + 343.
6. Köhler, Wolfgang. 1940. *Dynamics in Psychology.* New York. Pp. 1–158.
7. Lashley, K. S. 1923. The behavioristic interpretation of consciousness. *Psychol. Rev.*, **30**, 237–272; 329–353.
8. Lashley, K. S. 1930. Basic neural mechanisms in behavior. *Psychol. Rev.*, 37, 1–24.
9. Lashley, K. S. 1933. Integrative functions of the cerebral cortex. *Physiol. Rev.*, **13**, 1–42.
10. Lorente de Nò, R. 1934. Studies on the structure of the cerebral cortex. II. Continuation of the study of the ammonic system. *Jour. f. Psychol. u. Neurol.*, **46**, 113–177.
11. Lorente de Nò, R. 1938. Analysis of the activity of the chains of internuncial neurons. *Jour. of Neurophysiol.*, **1**, 207–244.
12. Sherrington, Sir Charles. 1941. *Man on his Nature.* Cambridge. Pp. 1–413.
13. Stevens, S. S. 1935. The operational definition of psychological concepts. *Psychol. Rev.*, 42, 517–527.
14. Thurstone, L. L. 1938. *Primary Mental Abilities.* Psychometric Monogr. No. 1. Chicago. Pp. ix + 121.
15. Wiesner, B. P., and Sheard, N. M. 1933. *Maternal Behavior in the Rat.* Edinburgh. Pp. xi + 245.
16. Woodworth, R. S. 1918. *Dynamic Psychology.* New York. Pp. 1–210.

THE GENESIS OF BEHAVIOR FORM IN FETUS AND INFANT

The Growth of the Mind from the Standpoint of Developmental Morphology

ARNOLD GESELL

Director of The Clinic of Child Development, Yale University

(*Read April 25, 1941, in Symposium on Recent Advances in Psychology*)

Form is the most ubiquitous of all scientific problems. It figures in one guise or another in the physical sciences, the life sciences, the social sciences. When Goethe coined the word *Morphologie* he was thinking particularly of the shapes of skulls and flowers. But now the concepts of morphology are entering into a vast variety of fields ranging from atomic structure to the celestial galaxies, from electro-encephalography to patterns of savage culture and designs for modern living. It has been well said, Form is the fundamental riddle!

A Morphological Approach

Surely we cannot escape the problem of form when we contemplate the rich pageantry of the ontogenetic patterning of human behavior. In its beginnings this behavior may seem aimless, inchoate. But if somewhat vaguely we call certain early movements "vermicular" we already acknowledge that they have form. If we call them "random movements" we commit a scientific error. The infant does not react at random. His seemingly random movements are channelized and shaped by his neuromotor equipment. When they are studied by morphographic methods, including the method of cinemanalysis, behold they have characteristic form! The first task of a genetic psychology is to identify the forms and the formation of the manifold reactions of fetus, neonate, and infant.

The morphogenesis of manifest behavior begins with the fetal stage when the organism is a scant inch in length. At 8 weeks, on oral stimulation, this Caesarian organism bends unilaterally; at 10 weeks it bends bilaterally, a kind of body swing;

at 11 weeks the fingers clasp; at 12 weeks the hands approximate as if to clap—a fetal prefigurement of the nursery game of pat-a-cake. At 14 weeks the fetus can sneer and swallow—and can wink even though the eyelids are still fused. At 16 weeks a single shallow gasp (*ex utero*); at 22 weeks, a grouped series of gasps; at 25 weeks, continuous pre-respiratory movements (*in utero*); at 26 weeks or later, a viable baby able to breathe.

A prematurely born infant is an anomalous, air-breathing fetus. Let us call him a *fetal-infant.* We have followed his development keenly because it reflects the behavior morphogenesis characteristic of gestation. We have made 80 behavior examinations of 37 fetal-infants with fetal or post-conception ages of from 28 to 40 weeks. Systematic visual, auditory, tactile and motor tests supplemented by naturalistic observations revealed significant maturity differences in the forms of behavior at advancing ages. These studies have shown that although the healthy fetal-infant makes a remarkably good adjustment to an abnormally untimely environment, the basic schedule of his behavior patterning is not upset. He remains faithful to his fetality.

In so doing he makes a fundamental contribution to the experimental morphology of behavior. Could any experiment more drastic be conceived to test the integrity of behavior form under the stress of a sudden, extreme and prolonged change of environment? If endowed with sufficient vitality the unblemished premature infant survives; but he gains no ontogenetic headstart. If he were born 8 weeks prematurely the general conformation of his behavior make-up at 28 weeks would be that of a normal 20-weeks old infant. This indicates that the primary and preponderant forces in the morphogenesis of early behavior are intrinsically determined.

Our normative studies of full term infants and comparative studies of infant twins by the method of co-twin control show that postnatal behavior growth is obedient to the same steady, continuous developmental mechanics which governs prenatal growth. There are no dramatic moments when the infant abruptly acquires a gift of tongue, or perception, emotion, judgment, reason, volition, imagination. All of his behavior forms are essentially products of growth. We cannot penetrate the innermost citadel of his subjective life; but if we closely observe how he comes into command of his eyes and hands, we shall have more than an inkling of how his psyche is built up into morphological structure.

Presently we shall show a film which pictures early forms of eye and hand behavior in the fetal-infant, in the full term neonate and in infants at advancing ages throughout the first year of life. We have used the cinema as our major research tool (1) to chart the patterning of infant behavior, (2) to identify and to investigate the patterns by cinemanalysis, which is essentially a morphological method. By this method which utilizes a hand operated projector on a work table, it is possible to freeze and to activate recorded movements at the observer's discretion (Fig. 1). Cinemanalysis is indeed a kind of dissection which anatomizes the patterns of behavior. It makes the recorded reactions as tangible as tissue. The histologist who with his microscope looks at a stained section of the thyroid gland or at an adeno-carcinoma transplanted into a frog's eye is no closer to reality than the analyzer who looks at the outlines of a behavior pattern which come through the lens of the motion picture projector, and are imaged on a tracing plate. Both observers are inspecting, analyzing, and, if need be, measuring the characteristics of form in terms of time and space.

(*By way of demonstration a film entitled* INFANT EYES AND HANDS *was shown. This film was compiled from original records in the photographic research library of the Yale Clinic of Child Development. The contents of the film are indicated by the accompanying illustrations. See also the Bibliographic Note, page 488.*)

THE MORPHOGENESIS OF EARLY BEHAVIOR

In order that we may get an impression of the total morphogenetic sweep the film begins with an animated rendering of the 4-week old human embryo when the prospective hands are a pair of buds just behind the gill arches (Fig. 2). Cells invade these diminutive stumps and cause them to elongate. Three segments emerge. The outer segment assumes the shape of a paddle. Five lobes appear on the edge of the paddle which transforms into a five-fingered hand presently equipped with muscles, tendons, volar pads and neurons. End organs by the thousands, like so many sentinels, establish themselves in the sensitive skin. Indelibly patterned friction ridges sculpture themselves in the palmar skin. Behavior patterns are laid down by similar morphogenesis. At 11 weeks the fingers flex—the beginning of grasp. At 18 weeks they grip. At 20 weeks the whole arm

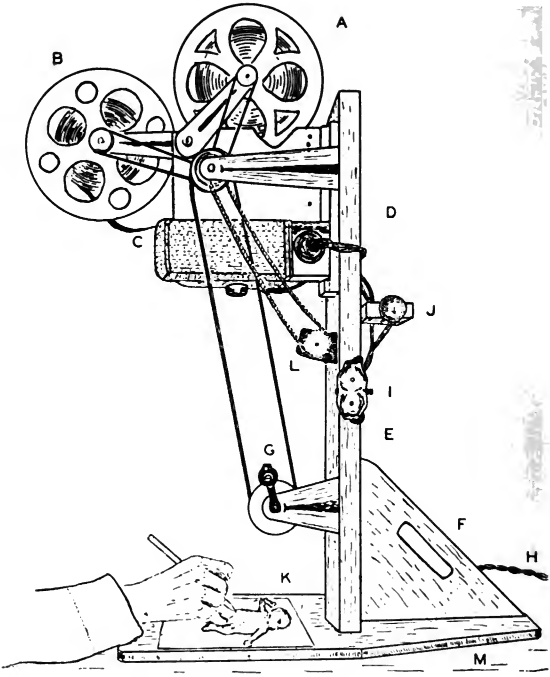


Fig. 1. Analytic viewer for the cinemanalysis of behavior patterns. The viewer consists essentially of a 16 mm. projector (*D*) mounted on a portable stand (*E*). The whole apparatus weighs approximately 16 pounds and can be readily carried by the handle (*F*). It may be conveniently placed on a work table or desk, the electric current being supplied through the extension cord (*H*), with a switch at (*I*). The projector may be driven by its electric motor, but ordinarily it is operated by the hand crank (*G*). The film (*C*) passes through the projector from a supply reel (*A*) to a takeup reel (*B*), each frame being registered by a frame counter (*L*). The projector, equipped with a 1 inch lens, throws an image on the tracing plate (*K*); the strength of the illumination is controlled by a rheostat (*J*).

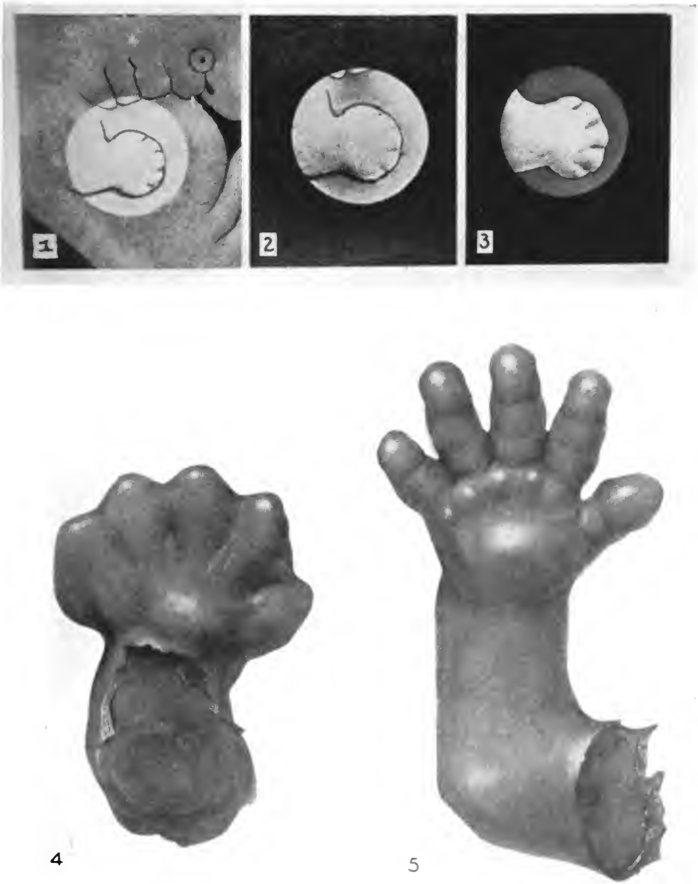


FIG. 2. The development of the fetal hand. 1, 2, 3, 4. Early limb bud stages. Embryo @ 17 mm. 5. Palmar aspect of right hand of a 24 mm. embryo, showing prominent apical and interdigital volar pads (after H. Cummins). Primarily these are *walking* pads. Secondarily they serve as *prehensory* and *tactile* pads.

extends laterally as the head turns to the same side—a head-hand coördination, the rudiment of the tonic-neck-"reflex".

This so-called t-n-r reflex is not a stereotyped entity but a growing thing which elaborates with the maturity and economy of the organism. It becomes well-defined in the fetal infant (Fig. 3). It is a normal characteristic of the full term neonate (Fig. 4). It persists for 16 weeks after birth. By 20 weeks the infant prefers a symmetric attitude, head in midline, hands approximating to the midline (Fig. 5).

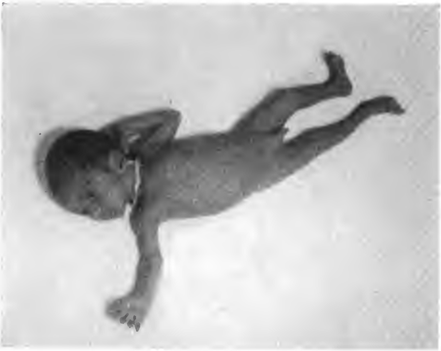


Fig. 3. The tonic-neck-reflex pattern in a fetal infant (Fetal age: 32 weeks). Note the head aversion, the extension of the faceward arm, the flexion of the opposite arm, simulating a fencing attitude.

The addiction of the young infant to t-n-r postures is both a symptom and a condition of his behavior growth. It represents a morphogenetic stage in which fundamental neurological coördinations are laid down to "form" the framework for later postural, manual, locomotor, and psycho-motor reactions. The t-n-r pattern complex is part of the ground-plan of the organism just as truly as the skeletal system, and it is shaped by comparable morphogenetic forces. The t-n-r attitude is a kind of matrix or scaffolding for the channelizing of oculo-motor postures, which lead to hand inspection and to a progressive coördination of the eye and hand movements culminating in prehension.

Just as the eyes of the fetus wink even before there is occasion to wink (the lids being fused!), so the primary correlations of eye and hand movements are relatively independent of experience. In this correlation the eyes take the lead. The infant lays hold of the world with his eyes before he grasps it with his hands.

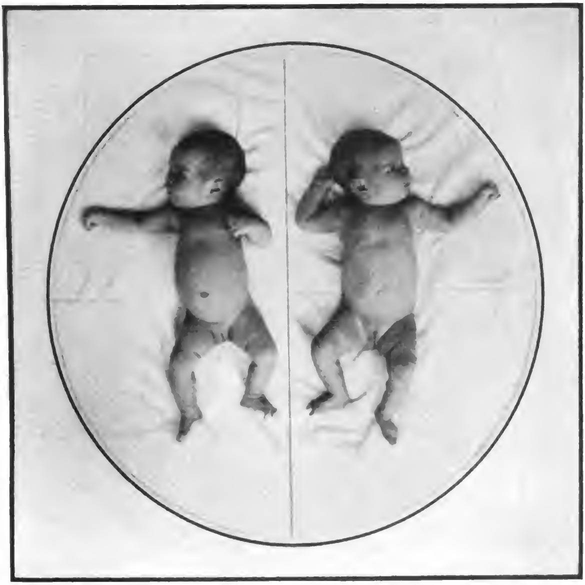


FIG. 4. The tonic-neck-reflex pattern in two full term infants, age 6 weeks. One infant was addicted to a right t-n-r, the other infant to a left t-n-r, a normal manifestation of asymmetry.

Conjugate deviation, *i.e.*, simultaneous deflection of the eyes in the same direction, enabling pursuit of a moving stimulus, is present to some degree in the first few weeks after birth. Likewise coördinate compensatory eye movements by which constant fixation of an unmoved stimulus object is maintained during rotation of the head. Incipient fixation of an approaching object is observable in the first day, sustained fixation of a near object in the first week, and perfectly sustained fixation for near and far objects at the end of the first month. At the end of the second month the ocular conquest of space is so well established that transfixed regard gives way to roving inspection: the infant explores the universe with binocular rather than monocular oculo-motor postures (Fig. 6).

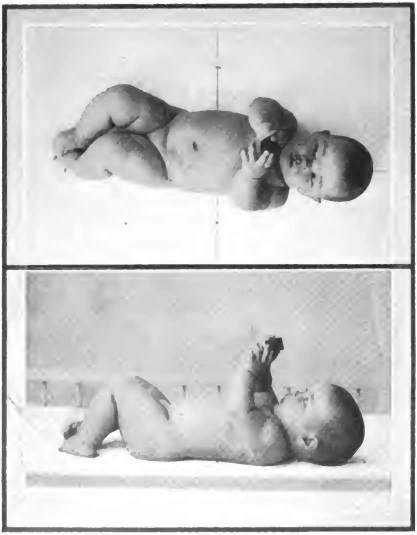


FIG. 5. Bilateral symmetry pattern in supine prehension. The asymmetry of the t-n-r gives way at about 20 weeks to symmetric attitudes. (Simultaneous photographs taken with one camera at the zenith and the other camera at the horizon of the photographic recording dome.)

These developmental preparations take place within a total action system and in close relation to postures of trunk and upper extremities. The t-n-r attitude leads to inspection of the extended hand and by gradual stages to prehensory approach, and to manipulation—unilateral, bilateral, alternating unilateral, and finally unidextrous. The early prehension of the infant is crude and precarious. He grasps palmarwise and massively. Later he grasps digitally with increasing reliance on the radial digits. At first the thumb is almost a useless member, but by nicely graded morphogenesis it moves from a position of prepivotal to pivoted adduction, through mesial to mesio-volar opposition to the index, and finally by circumduction to full opposition. Before the end of the first year the infant plucks a pellet with precise pincer prehension, and with almost adult skill (Fig. 7).

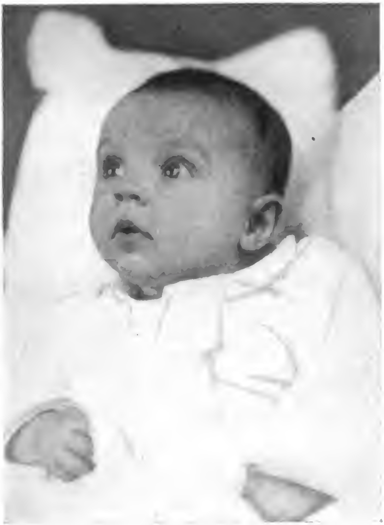


FIG. 6. Ocular fixation. This infant, age 16 weeks, has attained the fifth stage of oculo-motor control and is capable of both sustained and roving exploratory regard.

These transformations of prehensory pattern are profoundly determined. They are condensed versions of phyletic sequences from plantigrade to upright, from pawlike to digital, from pronate to oblique, and from ulnar to radial orientations. They are infused with "the mysticall Mathematicks" which makes the individual a representative of a species.

The precision of this morphogenetic mathematics is most beautifully demonstrated by the remarkable correspondences

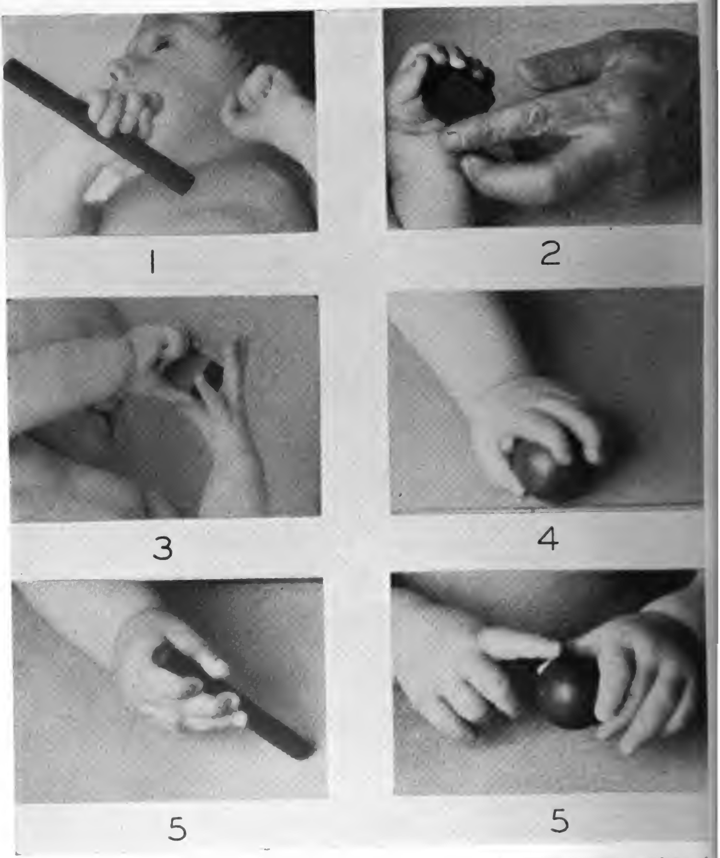


FIG. 7. The development of thumb opposition. Patterns of thumb posture typical of five maturity stages: 1, pre-pivotal adduction (0–16 weeks); 2, pivoted adduction (16–28 weeks); 3, mesial opposition (16–28 weeks); 4, mesio-volar opposition (24–36 weeks); 5, volar opposition (32–52 weeks).

displayed in the behavior development of highly identical twins. Twins T and C have been under our systematic observation for a period of fourteen years. Since early infancy there has been an amazing degree of concordance in the emergence and the configuration of their patterns of behavior. These patterns have been as much alike as the patterns of their fingerprints and of their hair whorls, even to the extent of exhibiting mirror imaging at times. Mirror-symmetry is reflected in the patterns of pellet prehension pictured in fig. 8. At 38 weeks the twins addressed a pellet with a fanlike spread of all the digits, the thumb extended at a right angle. At 40 weeks they made a raking approach; at 42 weeks they poked the pellet with the tip of the index. These changes in prehensory patterns took place synchronously and endogenously. Comparative studies of the behavior of Twins T and C by the method of co-twin control have shown that the endogenous mechanisms are so firmly entrenched that they can scarcely be transcended by training. All learning and conditioning are limited by a maturational substrate built up through the architectonics of intrinsic growth.

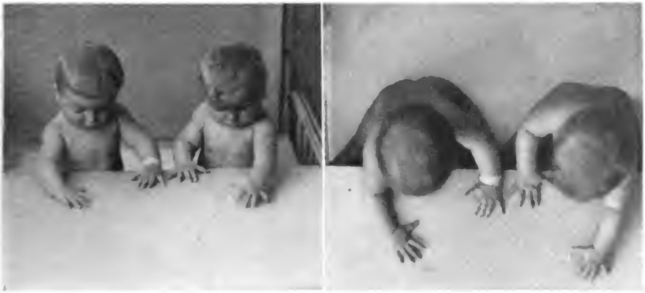


FIG. 8. Mirror-image symmetry in the prehensory patterns of one-egg twins. Simultaneous horizon and zenith views of twins T and C in the act of making prehensory approach upon a pair of pellets. Striking ontogenetic correspondences were observed in the development of their behavior patterns from infancy to adolescence.

The simultaneousness of behavior transformations observable in highly identical infant twins confirms a vague but pregnant passage from Shakespeare: "A man may prophesy with a near aim of the chance of things, as yet not come to life, which in their seeds and weak beginnings lie intreasured. Such things become the hatch and brood of time.''

The cinema hastens the hatching! Our film has compressed into a few minutes the twenty lunar months which lie between the limb bud stage and the stage of precise pincer prehension. By telescoping time in this manner we accentuate the spatial configuration and reconfiguration which characterize the processes of behavior growth. We see the succession of behavior forms as progressive reorientations in space, as a fluent morphosis in which every event is a pattern which issues into a contextual pattern, which engenders yet another pattern—a dynamic flow. It is as though on a vastly complex scale we were observing the growth of a crystal organizing its facets and its axes in a transparent solution.

Is the crystal only an analogy; or is it an expression of the very same ordering forces which govern the ontogenesis of behavior? This brings us back to our original problem, *the riddle of form.*

Behavior and the Hierarchical Continuity

The crystal is more than mere analogy if we accept the principle of hierarchical continuity, and if we do not exaggerate the increasingly meaningless distinction between the terms life and non-living. (Dr. William Stanley yesterday gave us good reason to reduce the distinction.) Is the tobacco virus alive? Where lies its virulent capacity of reproduction, if not in the inherent geometry of its huge molecules? Each molecule is a single virus. The molecules may cluster in crystalline needles. The crystal order in turn is a natural expression of the properties of matter. The biological order likewise. Organizing relations are found at every level—''at the molecular level and at the colloidal and paracrystalline level as well as at the anatomical level'' (Needham).

Behavior caps the hierarchical continuity. Behavior forms stand at the very summit of the biological order. But are they, therefore, necessarily unique? Are they not, in fact, simply the most subtle and exalted manifestations of anatomical organization? Until there is proof to the contrary we should regard them as members of a single, ordered continuum.

By the principle of hierarchical continuity there can be but one physiology of development. The growth of tissues, of organs, and of behavior is thus obedient to comparable if not identical laws of developmental morphology.

From the pattern of a protein molecule to the neurological engram of a behavior pattern is indeed a far cry. Yet science may some day bridge the distance. (Elaborate Professor H. S. Taylor's atom models; add to the snap-fastener side chains a multiple array of zipper devices under photo-electric control in a bio-electric field-system and you have a working model of the syntax of living behavior form!)[1]

Genetic psychology as well as chemistry must reckon with form concepts. Stereochemistry considers the structural configuration of the molecule in three-dimensional space. Genetics is envisaging the infinitesimal morphology of chromosomes and genes. The gene is pictured as an organic particle, which like the virus may be a single large molecule. It has the remarkable capacity of propagating itself and of reorganizing the surrounding molecules, a process well likened to a contagion of geometry. Wrinch conceives the chromosomes in terms of warp and woof—long, parallel filaments of protein molecules, each filament surrounded with ring-like molecules of nucleic acid. Peters postulates for every cell a cyto-skeleton of three types of protein, corresponding to the receptor, conductor, and central elements of a nervous system.

Bio-chemists speak as familiarly about the dorsal, ventral, facial, and left-handed and right-handed aspects of their molecules as we do of the corresponding aspects of a baby. If there is a uniting hierarchical bond between these remotely separated fields of discourse, it probably lies in the varied morphogenetic factors or "forces" which are now being accurately explored by exact and experimental methods. Among these forces are valences, surface films, metabolic gradients, bio-electric fields, evocators, autonomous induction, regional determination, competence, time correlation, polarity, symmetry, and crystalline protein fibers oriented into patterns by hormonic chemical substance.

MORPHOGENESIS BY MATURATION

In last analysis we must look to genes for the knowable origins of behavior form. The genes in their myriad multiplicity

---

[[1] Cf. Address by H. S. Taylor: "Large Molecules through Atomic Spectacles," April 24, 1941. To be published in PROC., Vol. 85. Ed.]

are like a leaven distributed in each and every growing cell of fetus and infant. They are the chemical agencies which participate in those obscure but precise events which create form, not only the self-sculpture of bones, but also of brains and of behavior. They engender the sequences of ontogenesis. They mould its manifestations. They justify the concept of maturation, as the primary determiner of human behavior, prior to experience. Experiential and environmental factors inflect and specify the details of configuration, but the primary, the provisional and the prospective components of patterns are intrinsic in origin. Maturation is the net sum of the gene effects, and as such is the basic determinant of behavior form.

There seems to be little scope for the old doctrines of undifferentiated plasticity and of random movements as explanations for the genesis of behavior forms. Plasticity can exist only in relation to concrete situations and specific conditions in a formed and forming organism. It is in no sense a homogeneous property mysteriously diffused throughout the organism. On the contrary it is always a partially defined potentiality appropriate to a given stage of maturity. Plasticity has shape.

Behavior patterns develop from the beginning by an expansion of a total reaction system which is essentially a dynamic complex of postural sets. The morphogenetic steps by which a pattern acquires full maturity are typically as follows:

0. *Pre-nascent Stage.*—Complete absence of the function which later is embodied in a behavior pattern.

1. *Nascent Stage.*—Imperfect, inadequate, sporadic manifestation of the function in loose and variable associations with several postural sets.

2. *Assimilative Stage.*—More positive performance of function which, however, is dependent upon particular postural sets, and accessory reinforcing postural attitudes.

3. *Coördinating Stage.*—Perfected performance limited to particular postural sets but with sloughing off of the accessory postural attitudes, previously necessary.

4. *Stage of Synergic Individuation.*—Independence from restricted postural sets; versatile performance smoothly synergized with numerous and varied postural sets.

Growing behavior therefore is never random, because it is circumscribed by the attained morphology of the organism as evidenced in its postural sets. The apparent randomness is not some wild, haphazard display of a principle of uncertainty, but it is an active, ordered variability of a growing "margin". In the language of experimental embryology, the margin may be conceived as a state of dis-equilibrium in a complex system of reactants, which reactants are in last analysis genes or gene products operating as morphogenetic agencies.

This does not solve our riddle of form, but it does help to clarify the unknown. The concept of maturation obliges us to regard conditioning and learning as ancillary rather than primary mechanisms. It invites us to investigate even the most recondite problems of mentation and of psychic constitution from the standpoint of a dynamic developmental morphology.

* * *

A Bibliographic Note

The modern interest in the developmental aspects of circumnatal behavior may be dated from the significant studies of Wilhelm Preyer who in 1885 published his famous *Specielle Physiologie des Embryo.* This work was recently translated by George E. Coghill and Wolfram K. Legner and appeared as a monograph of the Society for Research in Child Development under the title, *Embryonic Motility and Sensitivity.* Coghill's own life work has been devoted to correlated anatomical and physiological studies of the growth of the nervous system in amphibia. His monumental studies are summarized in three London lectures under the title, *Anatomy and the Problem of Behaviour,* a slender volume which bids fair to become a classic.

Minkowski in 1920 began a series of observations on the behavior of surgically removed fetuses. A summarizing contribution appeared under the title, *Neurobiologische Studien am menschlichen Foetus.* Davenport Hooker has made valuable extensions of these observations documented by motion picture records. His paper on "The Origin of the Grasping Movement in Man" is published in the *Proceedings of the American Philosophical Society.* An earlier paper appeared in the *Yale Journal of Biology and Medicine,* entitled, "Early Fetal Activity in Mammals." The presidential address delivered at the 48th Annual Meeting of the American Psychological Association at State College, Pennsylvania, in September 1940, bears the title, "The Experimental Embryology of Mind." This address by President Leonard Carmichael deals particularly with the sensory aspects of fetal behavior as experimentally observed in cat and guinea pig, affords a survey of related studies, and a bibliography of sixty-four titles (*Psychological Bulletin*, 1941).

A scientific journal entitled, *Growth*, was recently founded for the coördination of studies of development and increase as general properties in nature. This journal brings under one cover biometric, anatomical, biochemical, and also behavioral studies. The journal is symptomatic of the increasing correlation and integration of the life sciences. Joseph Needham erected a bibliographic landmark in 1931 with his three volumes with the innovational title, *Chemical Embryology*. They have already been measurably outdated by the extraordinary productiveness of research in the field of experimental embryology.

Our studies at the Yale Clinic of Child Development have been concerned with the systematic delineation of the forms and early growth of human behavior patterns, from the period of fetal infancy throughout the first five years of life. *An Atlas of Infant Behavior,* in two volumes, illustrated with 3,200 action photographs deals especially with the first year of life. A handbook entitled *The Psychology of Early Growth,* serves as a key to the Atlas. A collaborative volume by the staff of the Clinic surveys development of behavior in the preschool years, under the title, *The First Five Years of Life: A guide to the study of the preschool child.*

The photographic research library of the Clinic consists of an extensive catalogued collection of motion picture records which chart the developmental progressions of behavior patterns, in infancy and early childhood. Both normal and abnormal behavior has been recorded under standardized and naturalistic conditions. The patterns are subjected to detailed examination by the method of cinemanalysis. The special studies bearing on the subject matter of the present paper are listed in the references.

In his introductory remarks Mr. Roland S. Morris, as President of the American Philosophical Society, and as presiding officer at the special session devoted to psychology, suggested that practical applications as well as recent advances might be considered. It may be pointed out in a general way that all science which is fundamentally addressed to a better understanding of the processes of development and the mechanisms of behavior must have some implications for practical utilization. Increased knowledge will ultimately lead to a better control of human growth.

Developmental norms of child behavior find direct application in clinical medicine, particularly in the fields of pediatrics, neurology and psychiatry. Such norms are necessary for the timely detection of defects and deviations of infant development. Clinical methods and applications are presented in a current volume entitled, *Developmental Diagnosis: Normal and Abnormal Child Development.* From the standpoint of development morphology the phenomena of infant behavior must be approached with the same minute interest in structured form which the disciplines of embryology and anatomy demand.

REFERENCES

COGHILL, G. E. 1929. *Anatomy and the Problem of Behaviour.* Cambridge, England: University Press; New York: Macmillan, pp. xii, 113.

GESELL, A. *et al.* 1934. *An Atlas of Infant Behavior:* A systematic delineation of the forms and early growth of human behavior patterns. Volume I. Normative Series. Volume II. Naturalistic Series. New Haven: Yale University Press, pp. 922.

GESELL, A. 1935. Cinemanalysis: A method of behavior study. *J. Genet. Psychol.,* 47, 3–16.

—— AND HALVERSON, H. M. 1936. The development of thumb opposition in the human infant. *J. Genet. Psychol.,* **48**, 339–361.

—— AND THOMPSON, H. 1938. *The Psychology of Early Growth, Including Norms of Infant Behavior and a Method of Genetic Analysis.* New York: Macmillan, pp. ix, 290.

———— 1941. *Twins T and C from Infancy to Adolescence.* A biogenetic study of individual differences by the method of co-twin control. Genetic Psychology Monographs (in press).

GESELL, A. 1938. The tonic neck reflex in the human infant, *J. Pediatrics,* **13**, 4, 455–464.

—— *et al.* 1940. *The First Five Years of Life.* A guide to the study of the preschool child. New York: Harper & Bros., pp. xiii, 393.

—— AND AMATRUDA, C. S. 1941. *Developmental Diagnosis: Normal and Abnormal Child Development.* New York: Hoeber, Medical Book Dept. of Harper & Bros., pp. x, 430.

HOOKER, DAVENPORT. 1936. Early fetal activity in mammals. *Yale J. Biol & Med.,* **8**, 579–602.

—— 1938. The origin of the grasping movement in man. *Proc. Amer. Philos. Soc.,* **79**, 597–606.

LING, BING-CHUNG. 1941. A genetic study of sustained visual fixation and associated behavior in the human infant from birth to six months. *J. Genet. Psychol.* (in press).

MINKOWSKI, M. 1928. Neurobiologische Studien am menschlichen Foetus. *Abderhalden's Handb. d. biol. Arbeitsmethoden,* Lief. 253, S. 511–618. (This is the concluding and summarizing communication of a series begun in 1920 and consisting of ten papers.)

NEEDHAM, JOSEPH. 1931. *Chemical Embryology.* Cambridge, England: University Press, pp. xxi, 2021.

PREYER, W. 1937. Specielle Physiologie des Embryo (Leipzig, 1885). Translated by George E. Coghill and Wolfram K. Legner: Embryonic Motility and Sensitivity. *Monograph of the Society for Research in Child Development,* II, No. 6 (Ser. No. 13), pp. 115.

Yale Films of Child Development. Infant Eyes and Hands. 16 mm., silent, 400 feet, 1941. The Development of Thumb Opposition in the Human Infant. 16 mm., silent, 100 feet, 1936. The Tonic Neck Reflex (t-n-r) in the Human Infant. 16 mm., silent, 130 feet, 1938. The Growth of Infant Behavior: Early Stages. 16 mm., silent and sound, 400 feet, Erpi Classroom Films, Inc., 1934. The Growth of Infant Behavior: Later Stages. 16 mm., silent and sound, 400 feet, Erpi Classroom Films, Inc., 1934.

ON THE NATURE OF ASSOCIATIONS

WOLFGANG KÖHLER
Professor of Psychology, Swarthmore College

(Read April 25, 1941, in Symposium on Recent Advances in Psychology)

Among the functional principles of present-day psychology the concept of association is perhaps the most widely used. As a term, it is true, "conditioning" seems almost more popular among American psychologists. But then, as a principle of empirical connection, conditioning resembles association in so many points that the assumption of a basic factor common to both is fairly forced upon the theorist.

But what exactly do we mean when we speak of associations? In the early history of association theory it was customary to enumerate three classes of associations: the association by contiguity, *i.e.*, by neighborhood in space and time, the association by similarity, and the association by contrast. It is important to realize that only the first class actually involves an association, namely, the establishment and the persistence of a connection between items because these items have occurred together. What was called an association by similarity, for instance, was not assumed to presuppose such a connection. Rather, if an experience *A* had once occurred a subsequent similar experience *A′* was said to cause recall of the previous experience *A* merely because the two were sufficiently similar; in other words, without any pre-established connection of *A* and *A′*. Strictly speaking, therefore, the so-called association by similarity was not supposed to be an association at all; rather it was a particular form of *recall*. It is most unfortunate that in this manner the term association was used in two different meanings, for the most part with the connotation of a connection established by previous experience, but often also as a synonym for recall. The consequences of this ambiguous terminology have not yet been fully overcome in our theories of learning and of memory.

In the following pages a sharp distinction between association and recall will be particularly relevant because our discussion will deal with influences which varying degrees of similarity between given items exert on the strength of *associations.* Whenever we use this term we shall give it its proper meaning in which it refers to empirical connections. On the other hand, problems of recall, including recall by sheer similarity, will not be considered.

Are associations in this sense strictly speaking associations by mere contiguity? Such a term seems to imply that neighborhood in space and time is the essential factor which establishes connections between processes or between their neural traces. About twelve years ago, however, Professor Thorndike raised objections to this simple interpretation. An association, he said, will not be formed by mere neighborhood of the items in question; these items will be associated only if they are experienced in a relation of mutual *belonging.* At first sight the introduction of this concept seems to imply a fundamental change in association theory. In fact, certain experimental examples by which Thorndike supported his view are likely to suggest that the term belonging refers to a rational or sensible relationship. This, however, would be a misunderstanding of Thorndike's own intentions. He expressly warns his readers against this interpretation and states that any "this goes with that" appears quite sufficient as a realization of belonging. Thus Professor Thorndike assumes an attitude of extreme caution, a caution which, however, has its disadvantages. It will now be difficult to distinguish between instances of mere contiguity and others in which over and above contiguity "this goes with that."

Another attempt to modify the concept of contiguity has been made by Gestalt psychologists. In the development of association theory the connection between the associated items had gradually become an entirely different bond; in physiological terms, an anatomical pathway along which excitations travel with particular ease from the locus of one item to that of another. It was assumed that contiguity of items in space and time makes for better conduction along the fibres and through the synapses which lie between the anatomical loci of those items. Apparently the nature of the items themselves had little influence on their association. At least the relation which the characteristics of one item bear to the characteristics of the other seemed to play no part in the process.

It occurred to some of us that much more than this might be involved in an association. In primary experience, and particularly in perception, Gestalt psychologists lay great stress on what they call organization. The most elementary characteristic of organization consists in the fact that the contents of specific areas in space and those of specific stretches in time are experienced as unitary entities which appear as segregated from their general environment.

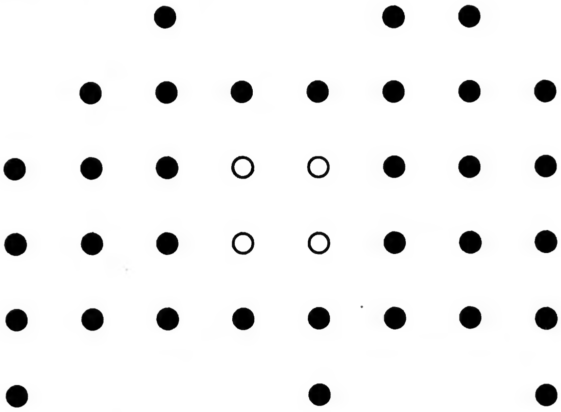


Fig. 1.

It will suffice if I give merely one example (Fig. 1). As an instance of organization this example may not be particularly impressive; but it refers directly to a peculiarity of organization with which we shall be concerned in a moment. In this pattern four circles are spontaneously united and singled out as the corners of a square. These four circles are outline-figures while the others are entirely black. Geometrically two outline-circles and two black ones would be just as good as corners of the square. But under ordinary conditions such an organization will not occur in the pattern of Fig. 1. In the formation of a perceptual group *like* items tend to unite, and where a new kind of item begins there the unitary group tends to end.[1] It is well known that *proximity* in space and time has the same effect as similarity or likeness. Other things being equal, items which are near one another unite most readily in one group, and this group tends to have its boundary where distances become greater.

---

[1] Red circles among many black circles or any analogous distinction would serve as well for a demonstration of organization by similarity.

Organization in this sense is not yet a generally accepted concept. To many it will appear mysterious until it has been interpreted in terms with which they are more familiar. Actually, no mystery is involved in this concept. To be sure, many details of organization need further elucidation, and the theory of organization is only now beginning to assume a more definite shape. The fundamental idea, however, which we have to use in this connection is simple enough: it is the concept of *interaction.* If under the influence of peripheral stimulation various processes develop in a given sector of the nervous system such events are not likely to remain separate and independent. They are likely to *interact.* Interaction among neural processes will tend to be specific and selective in the same sense as is interaction, say, in chemistry. Not every process will interact indifferently with every other process. As in all physical or chemical interaction, the specific properties of the interacting processes in their relations to one another will play a part in the determination of the result. It appears that in the case of perceptual organization *resemblance* is a condition which greatly favors interaction. Moreover, just as all interaction in physics is favored by shortness of distance so perceptual organization follows the principle of *proximity.* Obviously, the forces involved in an interaction are stronger when acting over short distances. This seems to hold in the nervous system as it does in inanimate nature.

I can now introduce my experimental problem. The theory of learning and memory must assume that the neural events which accompany primary experience do not entirely disappear when this experience vanishes. If no aftereffects, no *traces,* of these events were left in the nervous system, learning would be unable to influence future activities; in other words, memory in the widest sense would be impossible. Suppose now that items appear as parts of a unitary experience because interaction transforms their neural counterparts into dependent members of one coherent process. Under such circumstances this unitary process will be followed by a neural trace which has the same unitary character. What will happen if at a later time a part of this unitary trace is once more thrown into action? Because of the unitary character of the trace, this excitation will spread more easily within the trace than to other regions of the tissue. Thus the parts of the trace which are not directly excited will be indirectly excited. But this is just the physiological event which we suppose to happen in recall "by contiguity," *i.e.,* when two items *A* and *B* are associated, and when therefore re-excitation of *A* leads to re-appearance of *B*. Differently expressed, the fact that a unitary process forms a unitary trace is equivalent to what we call an association. From the present point of view association is therefore simply coherence within the unitary trace of a unitary experience. Since such coherence is the outcome of specific interaction or organization among the primary processes, the concept of association becomes a secondary concept. It can be derived from the concept of organization and from that of traces in general.

At first sight this reasoning may appear a bit vague. Actually, however, it forces us to draw certain conclusions which can be subjected to experimental test. The principle of such conclusions is extremely simple: If association is coherence within the unitary trace of an equally unitary or organized experience, then any principle of organization in primary experience must *ipso facto* be a principle of association, too. More specifically, whatever factors favor organization in primary experience must at the same time favor association, retention, and therefore recall.[2]

As yet I have mentioned two such factors: both proximity and similarity of items favor their organization in a unitary experience. Do the same factors actually favor association? As to the first principle, proximity, the answer is obvious. Proximity is practically a synonym for contiguity. In other words, if experiences are contiguous in the sense of association theory, interaction between their neural counterparts will be facilitated by the factor of proximity. Thus the law of association by contiguity holds because organization depends on proximity.

At this point, therefore, the main body of classical association theory seems simply to be absorbed by the organization theory of learning. And yet the transition from one view to the other implies much more than a mere change of terms. In organization theory proximity is essential because it favors *interaction.* This concept of interaction among items does not occur in association theory. And from this concept more follows than is expressed by the statement: contiguity is practically the same as proximity.

---

[2] In the case of recall this general rule is subject to a certain limitation. But as this limitation follows directly from the principle of organization itself, it constitutes no argument against the present theory. *Cf.* the author's *Gestalt Psychology* (1929), pp. 311 ff.

Nowhere is interaction independent of the characteristics of the interacting objects or processes. If these characteristics and their relation are changed the interaction, its modes, and its strength will be changed at the same time. Thus organization in primary experience depends not only on proximity but also on the degree of resemblance which the materials in question exhibit. If therefore our basic assumption about the nature of associations is right the likelihood of effective association, too, must depend upon the resemblance between the items in question. In experimental terms: If subjects are asked to learn pairs of items so that associations are formed, they should succeed more easily if the two members of the pairs belong to the same category than if these members have widely different characteristics.

During the past few months this conclusion has been subjected to a number of elementary tests. I wish to thank Miss J. Goldstein of the University of Pennsylvania for the skill and patience with which she carried out many experiments. Several undergraduates of Swarthmore College and one of Haverford were indefatigable in gathering further experimental evidence. Dr. H. Wallach of Swarthmore assisted me in devising the experimental procedures.

A first test had this form: To a group of six subjects (A) a series was presented which consisted of 6 *homogeneous* pairs of items, namely, of 2 pairs of nonsense syllables, 2 pairs of two-place numbers, and 2 pairs of nonsense figures. A second group (B), again of six subjects, was given a series of 6 *heterogeneous* pairs, a syllable with a number, a figure with a syllable, and the four remaining combinations. Thus both series contained 4 syllables, 4 figures, and 4 numbers; but in the first series these items were properly paired, in the second they were paired, so to speak, against their nature. Members of one pair were always simultaneously presented. From one subject to the next the sequence in the series was rotated by one step so that—with 6 pairs and 6 subjects—each pair occurred once in every possible serial position. Both groups were tested shortly after two presentations of their respective series. We used the procedure of paired associates, in which the subject is shown the first member of each pair and tries to recall its partner.

TABLE I

	Group *A*	Group *B*	Total	In %
Hom.	23 (36)	25 (36)	48 (72)	67
Het.	11 (36)	11 (36)	22 (72)	29

A fortnight later the two groups of subjects were given the same series; but subjects who had first learned the homogenous pairs had now to learn the heterogeneous pairs, and *vice versa.*

Table I contains the results of this experiment in terms of correct recall. The numbers in brackets indicate the frequencies of the test cases. In both groups, irrespective of the sequence of the tests, recall is more than twice as frequent when the pairs are homogeneous. Moreover (Table II) the same holds for the three materials taken separately, whether or not we fractionate for the two groups of subjects.[3] Even a severe statistician will

TABLE II

	Fig.		Syll.		Nu		Total
	A	*B*	*A*	*B*	*A*	*B*	
Hom.	7	9	7	7	9	9	48
Het.	5	4	2	3	4	4	22

---

[3] When results are computed for the three materials taken separately the case of heterogeneous pairs requires special consideration. Which are the heterogeneous instances, say, in the case of syllables, those in which syllables are followed by items of another class or those in which syllables as second members are paired with first members of another category? Our choice is not free for the following reason: Correct recall is clearly defined for syllables and for numbers; but it is less well defined for the figures, which the subjects *draw* as well as they can. As a criterion of correct recall I use in this case unambiguous reference of the drawing to the actually correct figure. This, however, is a fairly liberal criterion. Therefore, in order to make recall in the heterogeneous case comparable to that in the homogeneous case the results for pairs of figures have to be compared with heterogeneous cases in which the *second* member is a figure. Under these conditions the same criterion must also be applied to syllables and numbers, so that, for instance, a heterogeneous case under the category

admit that we can regard these results as reliable without much further computation. Our impression will be confirmed if we consider the results of the individual subjects who, we remember, were all tested in both constellations. Now we are fractionating for extremely small figures: in a given series a subject cannot recall more than maximally 6 items. And yet among 12 comparisons of the two results of 12 subjects there is only one which does not show a difference in favor of the homogeneous pairs (Table III).[4]

TABLE III

Group *A*		Group *B*	
Hom.	Het.	Hom.	Het.
3	4	5	1
6	4	1	0
6	2	2	0
2	0	6	3
3	1	6	5
3	0	5	2
23	11	25	11

In a second experiment we went one step farther. Individual items may offer varying difficulties to learning and recall. In the next experiment, therefore, the second items of the pairs were the same in both experimental constellations. Otherwise the construction of the series remained unchanged. Thirteen subjects were tested with the homogeneous and also 13 subjects with the heterogeneous pairs. The result is shown in Table IV. Independently of the material used, the homogeneous pairs give clearly better recall than the heterogeneous pairs, in spite of the

---

syllables may be either a pair: number-syllable, or a pair: figure-syllable. Of course, results *may* also be compared when the other possible classification of heterogeneous pairs is used. It is then found that, practically, the choice matters less than one might expect.

[4] Incidentally, this one subject reported spontaneously that he tended to make heterogeneous pairs more homogeneous, for instance, by discovering a resemblance with a number in a figure which was the partner of an actual number.

It may be just as important for the interpretation of these experiments as it is surprising that, generally speaking, our subjects did not realize what principle was followed in the construction of the series. Apparently the effort to learn all but prevented such considerations. One subject, a professor of psychology, remarked after his second test that this time learning had been clearly more difficult, but that he did not know the reason. As a matter of fact, in his second test he had been given the series of heterogeneous pairs.

fact that the items to be recalled are exactly the same. As the objection might be raised that the two constellations were learned by different subjects of possibly different average memory, 6 subjects of each group were weeks later asked to learn series of the type which had originally been learned by subjects of the other group. The second items of the pairs were again identical in the two series. The next table (V) shows how these 12 subjects fared in their two tests. Group *A* consists of the 6 subjects who were first tested with the series of homogeneous pairs and afterwards with the heterogeneous pairs; in Group *B* we have the 6 subjects for whom this sequence was reversed. It will be seen that results are invariably superior when homogeneous pairs are learned and tested, irrespective of subjects, of material, and of the sequence of the tests. On the average, association is twice as effective in homogeneous pairs of the present kind as it is in heterogeneous pairs.[5] The reliability of this result may be judged from an examination of smallest samples: The next table (VI) gives the scores of our 12 individual subjects in their two tests. No subject shows better recall in the hetero-

TABLE IV

	Fig.	Syll.	Nu	Total	In %
Hom.	19	19	17	55 (78)	71
Het.	12	8	7	27 (78)	35

TABLE V

Group		Fig.	Syll.	Nu	Total	In %
A	Hom.	9	10	8	27 (36)	
	Het.	2	5	6	13 (36)	
B	Hom.	7	8	11	26 (36)	
	Het.	5	5	4	14 (36)	
A + *B*	Hom.	16	18	19	53 (72)	74
	Het.	7	10	10	27 (72)	38

---

[5] In Table I and Table II which refer to analogous series results are in practically the same ratio as in Table V.

geneous series; with two subjects results are the same in both constellations; the remaining 10 individuals recall more in the series of homogeneous pairs than in the heterogeneous constellation.

Up to this point homogeneous and heterogeneous pairs have been tested in separate series. But our problem refers to *individual pairs* and to the degree of resemblance between their members. Consequently we ought to be able to demonstrate the same effect of similarity upon association when heterogeneous and homogeneous pairs are combined within one series. I need not describe in detail how the series was constructed in which this postulate was examined. Only figures and syllables were now used as crucial materials. Any item which some subjects had to recall in a homogeneous pair was given to the same number of subjects as the second item of a heterogeneous pair. Otherwise the principle of experimentation remained the same. No less than 60 subjects went through this test. Table VII contains the results. They prove that in a series which contains both homogeneous and heterogeneous pairs our rule remains valid: For homogeneous pairs scores are consistently higher than they are for heterogeneous pairs.

TABLE VI

Group *A*		Group *B*	
Hom.	Het.	Hom.	Het.
5	1	4	1
4	1	3	2
5	2	4	4
4	4	5	2
4	2	4	3
5	3	6	2
27	13	26	14

TABLE VII

	Fig.	Syll.	Total	In %
Hom.	41	38	79 (120)	66
Het.	16	21	37 (120)	31

I have been asked whether, even if the facts be granted, defenders of the traditional association concept need accept my interpretation of these facts. Perhaps they could argue that syllables go with other syllables in speech, writing, and reading, and that, similarly, numbers go with numbers in daily experience. Thus associations by contiguity are established long before a person becomes a subject in such experiments. These connections may favor the particular associations in homogeneous pairs, which are being formed in the present experiments.

I am not convinced by this argument. In the first place it will be difficult to apply the same reasoning to the case of our nonsense *figures.* In the second place previous associations of this kind will constitute vague reproduction tendencies in many different directions, and in a given case such a cluster of reproduction tendencies will not simply favor one specific association between items of the same class. Rather previous associations will exert inhibiting just as well as facilitating influences. In fact, if we had found precisely the opposite results of those which we have actually found, an explanation in terms of inhibitions by pre-established associations would have been equally plausible. And I am almost convinced that such an interpretation would actually have been proposed.

It seems preferable, however, to avoid such abstract discussions and to devise experimental tests which are less subject to similar arguments. Two such experiments will now be described. They are not yet completed but seem to corroborate our point of view.

In a first set-up *identical* heterogeneous pairs are given in two different forms of presentation. In both constellations a syllable is paired with a figure, a figure with a number, and so forth, so that in 6 pairs all possible combinations are exhausted. In one series all these items are of medium size, and all have the same red color on white ground; in the second series precisely the same items are shown, but now the first members are black and three times as large, the second members red and smaller just as in the first series. We remember that such factors influence perceptual organization. Do they influence association in the same fashion?

In the test the subjects are not asked to recall the color and the size of the second item; the test refers solely to the shape of the figures, to the numbers and the syllables as such. Our problem is whether nevertheless the greater ease of perceptual organization in the first series causes better association and on this basis better recall. Table VIII shows the results. In this table the terms "Homogeneous" and "Heterogeneous" refer to equality and to difference as to color and size.

TABLE VIII

	Fig.	Syll.	Nu	Total	In %
Hom.	10 (24)	8 (24)	11 (24)	29 (72)	40
Het.	5 (24)	3 (24)	8 (24)	16 (72)	22

Similar results were obtained when in a further homogeneous series—of again the same items—all items were black and had the larger size, while in a corresponding heterogeneous series the first members of all pairs were red and small, the second members black and large. If the totals of these two series are combined with those of Table VIII it is found that in homogeneous pairs (24 subjects) recall was correct in

64 out of 144 cases,

and that in heterogeneous pairs (again 24 subjects) the result was

40 out of 144 cases.

I do not contend that these data are conclusive. None the less there does not seem to be much doubt as to the final outcome when later the same experiment will be repeated with an improved technique.

In a last experiment an entirely different procedure was used. Primary organization depends not merely upon proximity and resemblance but also upon other important factors. Among these *"fitting"* is of particular interest. I will explain what is meant by this term. Organization is a form of interaction, in our present connection an interaction of processes in the nervous system. When such an interaction is completed a situation of equilibrium will be attained. Now, in physics we have a simple rule about the nature of equilibria, a rule which was independently established by three physicists: E. Mach, P. Curie, and W. Voigt. They observed that in a state of equilibrium processes—or materials—tend to assume the most even and regular distributions of which they are capable under given conditions. Moreover, Mach showed that this must be the case, because in a regular distribution the vectors of a physical situation will balance one another more fully than they do in less regular distributions. But equilibria depend upon the given conditions under which they develop. According to these conditions a high degree of regularity may be attained in one case of equilibrium, a lower degree in a second case. And an equilibrium will be the more complete or stable the higher the degree of regularity by which it is characterized. If we apply this

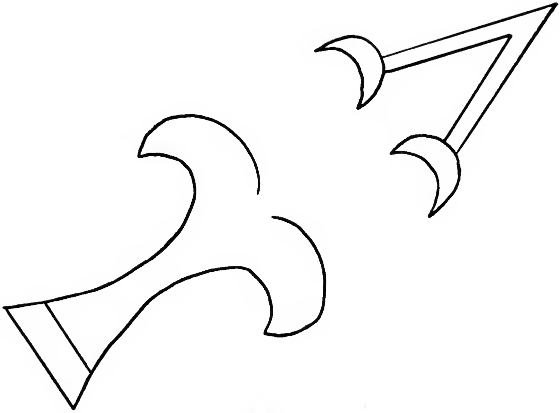


Fig. 2*a*.

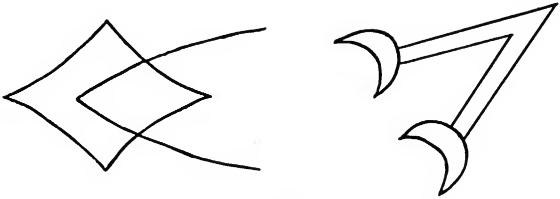


Fig. 2*b*.

to interaction in the nervous system it follows that regular organizations are the most stable products of such interaction. Suppose now that as external conditions we give in one case a pair of figures which form a regular total pattern, while in another case we combine two figures which represent a more arbitrary, a less *fitting* combination. Under these circumstances the first organization should be preferred in the sense that, once established, it is more stable. Facts of primary organization do exhibit this trend. From the viewpoint of this paper the same should hold for associations: Figures which fit one another in that they give a regular pair should yield better associations than do others which constitute less fitting pairs.

Preliminary experiments seem to show that this expectation is justified. Only *figures* were now used as items of our pairs. In one constellation these figures were combined into fitting pairs, in the other constellation precisely the same items appeared in more accidental and irregular combinations (*cf.* the examples in Fig. 2*a* and *b*). We found that second members of fitting pairs were more than twice as often recalled as were the same figures in irregular combinations. The difference was statistically reliable. A detailed report will be given as soon as more systematic tests have been completed.

To summarize: It seems no longer probable that association is an indifferent bond between merely contiguous items. Our evidence tends to support the view that associations are aftereffects of specific organization or interaction.

If future experiments should make this evidence conclusive our neurological conceptions would be confronted with a serious problem. Can we give adequate account of interaction in the nervous system if we think merely in terms of action potentials which travel along individual nerve fibres? We may gradually be forced to admit that, at least in ganglionic tissue, a more continuous type of function, a field function, is added to conduction along single fibres.

REFERENCES

KOFFKA, K. *Zur Analyse der Vorstellungen und ihrer Gesetze.* Leipzig, 1912.
KÖHLER, W. *Gestalt Psychology.* New York, 1929.
THORNDIKE, E. L. *The Fundamentals of Learning.* New York, 1932.

MENTAL ABILITIES

EDWARD L. THORNDIKE

Professor of Educational Psychology, Teachers College, Columbia University

(*Read April 25, 1941, in Symposium on Recent Advances in Psychology*)

Most of us use freely such terms as musical ability, mathematical ability, a good memory, a superior faculty of imagination, inventiveness, inability to control one's temper, inability to concentrate, poor judgment, lack of reasoning power, inferior intelligence, and lack of mental balance. We have more or less definite notions of what we mean by these terms and could point out persons who exemplified them. We could grade a hundred persons whom we knew intimately on a scale from low to high, or little or much, or weak to strong, for any of these abilities, though with some doubt and questioning. Popular notions of these mental abilities are, however, extremely vague and variable in comparison with popular notions of length, area, volume, weight, density, temperature, and many other physical facts.

Psychology has advanced beyond popular knowledge by constructing means of measuring these abilities, or, more accurately, means of measuring various behaviors which are symptomatic of these abilities. These mental meters, commonly called tests, began with such simple matters as the speed and accuracy of finding and canceling the A's on a page of mixed capital letters, or the number of digits that a person can remember correctly from a single hearing. The former was called a test of attention, and the latter a test of memory span.

At first psychology accepted the popular view that attention, memory, imagination, reasoning and the like were fundamental and unitary faculties or powers of the mind. If you had a notably superior memory it would be equally superior for words, numbers, faces, localities, etc. If your faculty of attention was weak, you would be unable to concentrate well on lessons, stories, games and all else. Consequently, it was expected that a dozen or score of rather simple tests would reveal the fundamental abilities of a person and measure them, at least, roughly.

Two lines of experimentation, begun about forty years ago, caused the abandonment of this expectation. The first studied the training of abilities and found such facts as the following: If a person practices finding and checking A's he gains notably in speed and accuracy, but his ability to find and check B's shows much less improvement, and his ability to find and check words containing two given letters, such as e and s, or n and o, or i and p, shows little or none. If a person is trained at memorizing series of digits or series of nonsense syllables, he may improve notably in the amount that he remembers from, say, ten minutes of study, but when he transfers his efforts to memorizing shapes, or poems, or passages of music, he shows very little improvement over the ability he had before the training. What has been improved by the training is not a general power of attention, or a general faculty of memory that operates regardless of what is to be memorized.

The second line of experimentation studied the correlations or covariances of different manifestations of an ability and found such facts as the following: If a thousand twelve-year-old boys are tested in respect of the ability to add integers, the ability to add fractions, the ability to divide integers by integers, and the ability to divide fractions by fractions, the correlations or covariances of these sub-abilities are far from perfect. The boy who is ablest in adding integers will not be the ablest in the other three sub-abilities. A boy's rank among the thousand will not be the same in all four, but will vary rather widely. Mathematical ability is not one thing, but a complex of many.

If the thousand boys take a number of tests chosen to measure intelligence, their scores on the tests that concern words will not correlate perfectly with their scores in the tests that concern pictures, or space relations, or numbers. Intelligence is not a unitary ability that operates regardless of the data on which it operates.

Mental abilities are not an orderly retinue of a few easily defined and unitary faculties or powers, somewhat like the chemical elements, for each of which a mental meter or test can be found by sufficient labor and ingenuity. A mental ability is a probability that certain situations will evoke certain responses, that certain tasks can be achieved, that certain mental products can be produced by the possessor of the ability. It is defined by the situations, responses, products, and tasks, not by some inner essence.

Except for certain powers which depend upon definite features of the eyes, ears, or other sense organs, intelligence is one of the most definite and unitary abilities that have been studied; and certain plausible hypotheses have been set up concerning a simple biological causation of it. But even it is regarded by prudent psychologists as a probability that certain mental achievements will be achieved, and is defined by these achievements.

Hence we find the most expert measurers of intelligence stating that an intelligence-test score measures only the ability in that test and whatever correlates perfectly with that score. Hence we find that an intelligence meter that is perhaps the most dependable yet devised is not called a test of intelligence, but only a test of intelligence CAVD, that is of sentence-completion, arithmetical problems, vocabulary, and the comprehension of directions or paragraphs. These facts seem like confessions of weakness in the science of psychology, but are really signs of health and strength.

Three or four hours testing of a person with suitable intelligence tests or meters gives a score which is highly indicative of how well he will do in school, how likely he is to escape confinement in an institution for the feeble-minded, and how well he will understand the sermons he hears, the policies for which he votes, and the like. It is usefully indicative of his fitness to make wise decisions as a parent, neighbor, and citizen. Such a score obtained at age eight to eighteen is probably equal in value to a careful physical examination at that age by an expert physician. It is a valuable supplement to the ratings by parents, teachers and others who have long and intimate acquaintance with the person.

It should, however, be noted that the ability measured by such tests is perhaps better named abstract intelligence, or intelligence operating with symbols. Intelligence operating with actual individual things—machines, storms, traffic jams, stoves, cakes, pies, etc., etc.—may differ considerably from the intelligence that determines scores in Stanford-Binet or CAVD tests. So also may intelligence operating with actual people,—as in the managing, persuading and comforting of a politician, foreman, salesman, or nurse.

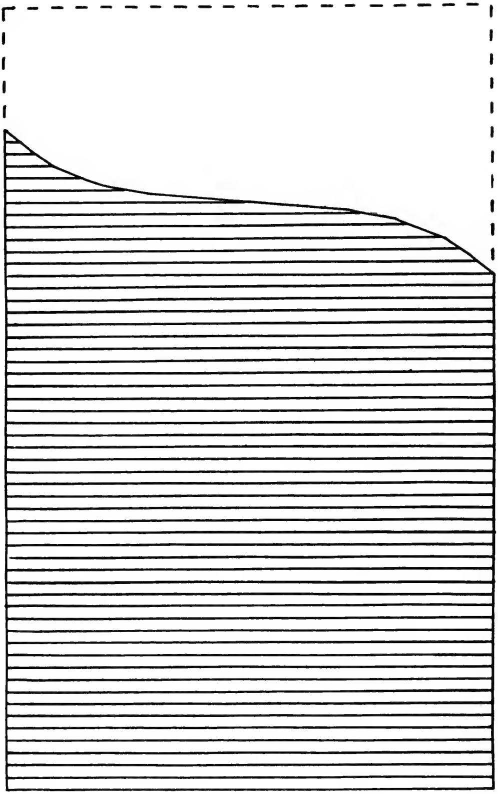


Fig. 1.

Our best present definitions and measurements of mental abilities are inventories of a person's actual or possible responses to certain tasks. The single score he obtains is some sort of an expression of the quantity and quality of his responses. It is customary to select or arrange the tasks so that each response will be completely right or completely wrong. The inventory can then be fully expressed by listing the tasks according to their difficulty and stating the percentage of successes at each level of difficulty. This concept of difficulty is of so great importance that I shall ask you to note a few illustrations of it. To give the opposite of *unless* is the same sort of task as to give the opposite of *up,* but is more difficult intellectually in the sense that fewer persons can do it and that those who can will average higher in intelligence by any reasonable measure of it. Similarly for supplying words to complete these sentences: A dog has four . . .; The relation between a number and its reciprocal is . . . . Similarly for answering these questions: What number is the sum of six and fourteen? What number is as much greater than half of six dozen as it is less than three-fifths of twenty dozen. In the case of musical ability it is more difficult to play a Beethoven symphony than to play chopsticks or Yankee Doodle, in the sense that fewer persons can do it and that those who can will average higher in musical ability by any reasonable measure of it.

Measurements of a person having much and a person having little of a mental ability which is an ability to do tasks alike in nature but differing in difficulty, are presented in Fig. 1 and Fig. 2.

The height of an ability, that is, the difficulty of the tasks the person can master, is more important than its width, because greater width can be obtained from more work by the person or from using more persons, but nothing can substitute for deficiency in height.

The scientific study of mental abilities early discovered that correlation, not compensation, was the rule. Nature does not balance superior reasoning power by inferior imagination, or superiority in learning by inferiority in remembering, or genius in music by idiocy in business. And this fact of the positive affiliations of estimable abilities has been confirmed again and again by later researches. This fact has very important consequences for man's life and work, mostly beneficent. But as far as concerns the exact description and measurement of abilities it is a nuisance. Psychology wishes to know each ability pure and undefiled, and to measure all of it and nothing but it. For example, it wishes three tests or meters of ability to appreciate color, rhythm and musical harmony. But it finds that in spite of its utmost ingenuity and effort, each test measures all three

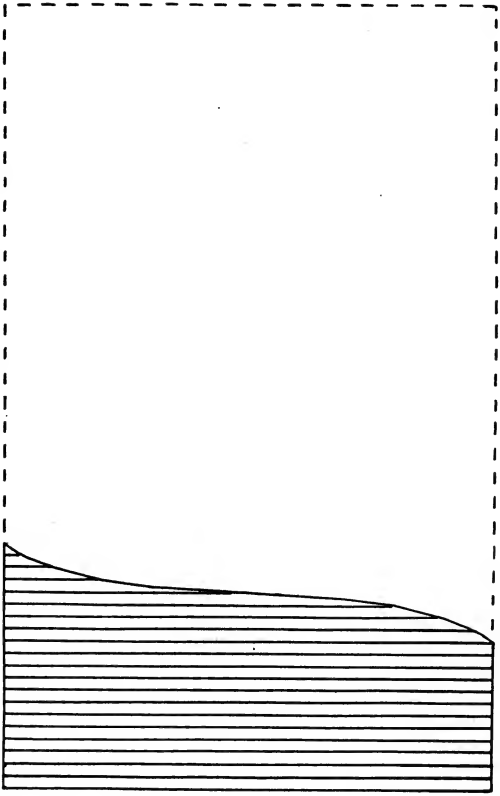


Fig. 2.

to some extent and also various potencies of the sense-organs, and also the person's general intelligence, so-called. Each score is contaminated, so to speak, by including more or less influence from other abilities. Each ability, as we actually measure it, is a mixture of abilities. We do not attain purity.

For various reasons psychologists would like to be able to replace a set of such impure abilities with positive intercorrela-

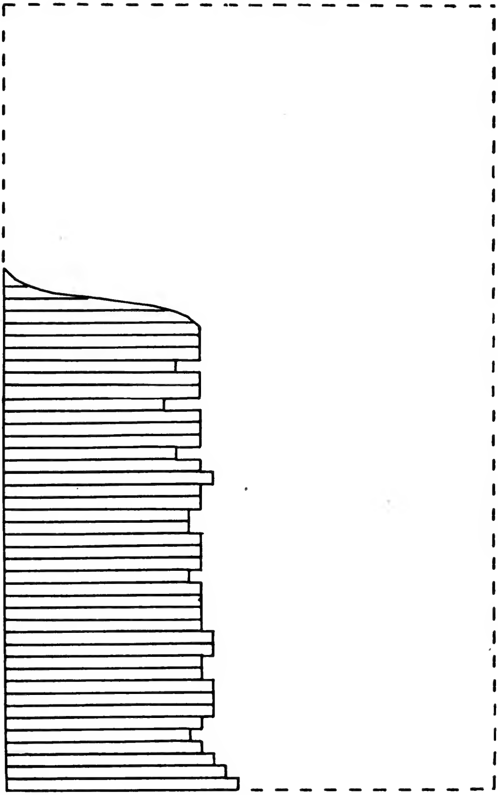


FIG. 3.

FIGS. 1, 2, AND 3. The entire area of the rectangle represents, in each case, a sampling of *n* tasks at each of 64 levels of difficulty for the ability in question, from a very easy task at the bottom to very hard tasks at the top of the rectangle. The shaded area represents the tasks accomplished successfully. The white area represents the tasks on which the person failed. Fig. 1 represents a person with a high ability. Fig. 2 represents a person with a low ability. Fig. 3 represents a person who had about the same number of successes as the person of Fig. 2 had, but who succeeded with much harder tasks, failing on many easy tasks. Cases like Fig. 3 occur rarely or never if the test is really of progressive levels of the same ability.

tions by a set of abilities that are pure and unitary or independent in the sense of having zero correlations one with another. And some psychologists cherish the hope that such pure abilities would prove to be more fundamental and more revealing of the mind's structure and of its physiological bases than the abilities found by direct observation and tests.

The methods of factorial analysis developed by Kelley, Hotelling, Thurstone and others succeed in discovering components or factors which would correlate zero one with another if we could get measurements of them in real people, and which are defined to a certain extent by their relations to measurable abilities. Work along this line has been carried on vigorously during the past five years, especially by Thurstone and his pupils. Few psychologists have mastered and used the methods of factor analysis, and those who have vary widely in their opinions of the results. Some expect that the factors discovered will inaugurate great advances in knowledge about, and control of, human abilities; some think that these factors are unrealizable abstractions, like an animal defined as 40 per cent man, 30 per cent turtle, 20 per cent shark and 10 per cent earthworm. The tests of the work will come when it progresses to the measurement of these components or factors, to using them for explanation and prediction, and to relating them to biological facts about the brain. This third test, however, is yet to be met not only by factor analysis but by the great bulk of psychological work on abilities as well.

It is fitting to note at this point that the important work of Spearman and his pupils, begun in 1904, was a stimulus to the recent work of factor analysis, though Spearman's main interest was in arguing that a group of abilities could, in the case of certain important groups, be represented by one factor common to all and many specific factors none of which was present in more than one of the abilities of the group.

Mental abilities are functions of the neurones or nerve cells that constitute the brain, sense organs and other parts of the nervous system, and psychology will not be satisfied until it succeeds in relating mental abilities to their neural cases, bases, parallels, or counterparts. Not much is known at present, and almost all of that concerns sensory abilities. The ability to see starts with specialized cells in the retina, the rods and cones.

The cones give sensations of color; the rods do not. The ability to see in bright light is different from the ability to see in dim light, and the rods play an important role in the latter. The two right-hand halves of the two eyes are put into intimate relation within the brain (and similarly for the two left-hand halves). Beginning with such simple facts as these and continuing to Dr. Selig Hecht's recent discoveries concerning the sensitivity of the visual purple, a great deal is known about the ability to see. But very little is known about the neural basis of the ability to draw or paint, or even the ability to have lively visual images, pictures before the mind's eye. Much is known also about the ability to hear, and psychologists are proud of Dr. Wever's share in this. But very little is known about the neural bases of the ability to understand language or to play the violin.

Relatively little is known, compared to what we would like to know, about the billions of neurones that compose the brain. We know that neurones conduct and a good deal about how they conduct. We have recently learned that neurones may act as glands secreting substances which may activate neurones which the secretion reaches. We know that the topography of the brain is influential on its action, certain functions being localized in certain systems of neurones. We know that the neurones in idiots' brains are relatively simple and coarse in structure. But if we had the brains of Newton, Beethoven, Darwin, Shakespeare and Bismarck all sectioned for examination, our best neurologists probably could not tell which had great mathematical ability, which had great musical ability, etc., or which knew English and which knew German. We are not even sure that we know where to look, or what to look for, in the search for the neural bases of such mental abilities.

The neurologists of two generations ago thought that they did know. They assumed that the brain consisted of "centers" each of which was of prime importance for some ability or abilities, and "pathways" from sense organ to center, center to center, and center to muscles and glands. A well developed "center" implied a high degree of the ability and was something like a combination of a factory and a storehouse. The so-called cell-body, that is, the thickened part of a neurone containing its nucleus, was supposed to be the main agent of neurones in causing abilities.

We now know that a brain is not much like a system of factories and storehouses interconnected by railroad tracks. It is more like a telephone and telegraph system with the added feature of more or less permanent modification by all the connections made and messages sent. It may include an elaborate system of resonators. What corresponds to an ability is more like a system of connections and modifications involving all army or navy activities than like a regiment of cell-bodies. The cell bodies indeed may only nourish and care for the conducting parts of the neurones. What the pattern of the neurones' actions is may often be more important than which particular neurones they are.

Localization of abilities in "centers" had been discarded a generation ago, and replaced by localization in groups of conductors and synapses with diffusion to other groups and convergence to common final paths. The localization was fairly complex for even such simple abilities as to scratch an itching spot, or clench the hand, or close the eyes, at will; and was increasingly complex for broader abilities. This doctrine of localization was attacked as too inflexible by a psychologist, Franz, who showed that the word done by one system of neurones could, to an unexpected degree, be taken over by another; and another psychologist, Lashley, has continued and greatly widened the attack in a brilliant series of experiments in which various parts of the brain of the white rat are extirpated and the resulting changes in the rat's abilities are observed and measured.

I lack time to describe Lashley's facts and inferences, and lack both time and ability to appraise them. But I think a consensus of competent students of the brain and mind would still hold that localization has real and great influence and that what an ability is depends in part upon which neurones are acting, but would admit that the localization within the cerebrum is far more complex and more indeterminate than had been supposed, and that it may be supplemented by selective sensitivities or resonances. Lashley's startling doctrine that the amount of cerebrum that is cut out rather than which part is cut out is often decisive for a rat's ability to learn may not be true for man, but it can reinforce a useful lesson—that in man every ability normally operates under an overhead control by his past development and present status and purposes. The man himself cooperates in most of his abilities.

I hope that my account of mental abilities has satisfied you. It will not satisfy my psychological colleagues. They will protest "Why did you omit all reference to the theories of the neural basis of intelligence that have been set forth by psychologists, including your own? Why did you say nothing about the differences between the sexes and between races, and the great differences between individuals? Why did you not evaluate the contributions of heredity and environment to various representative abilities? Why did you not trace the course of mental abilities from birth to maturity and on to old age? Why did you not report the careful work that has been done on certain special abilities, such as the ability to read and the ability to learn to sing?" Psychology has been active in all these fields and my only excuse for not including them is that if I had, this report would have been over twice as long.

PSYCHOANALYSIS AND SCIENTIFIC METHOD

CARNEY LANDIS

Principal Research Psychologist, New York State Psychiatric Institute;
Associate Professor of Psychology, Columbia University

(Read April 25, 1941, in Symposium on Recent Advances in Psychology)

The average experimental psychologist is often embarrassed by the supposition on the part of his scientific colleagues that psychoanalysis is part and parcel of psychology. To most experimental psychologists psychoanalysis is a source of generalized annoyance somewhat similar to that experienced by the astronomer when reminded of the subject of astrology. Again the professional psychologist listening to certain of his colleagues in psychiatry, neurology, anthropology or sociology, is all too frequently bewildered and amazed at the calm acceptance which his scientific brethren extend to psychoanalytic concepts and explanations. Our academic psychologist is astonished by essays written by his friends in the English or Greek Department wherein psychoanalytic terms are freely used and psychoanalytic theories expounded as fact. In the daily press he finds leading articles by columnists who couch their explanation of world events in psychoanalytic terms. His children come home from college talking about narcissism, repression and libido in the same matter of fact way in which they discuss democracy, vocational guidance, and religious freedom.

In brief, it seems to the psychologist that everyone is assuming that psychology is psychoanalysis and that the general public automatically assumes that any psychologist is also a psychoanalyst.

The professional psychologist is an experimentalist, whose research is designed in terms of the scientific method. Hence, in his approach to psychoanalysis he should make use of the same scientific methods which he applies when he studies learning, personality development, or perceptual illusions. Unfortunately, there seem to be only two avenues open for the investigation of the subject matter of psychoanalysis—reading what the professional psychoanalysts have written about it, or undergoing an analysis oneself. Most psychologists find the psychoanalytic literature inadequate from the standpoint of scientific reporting. The terminology is loose and ill-defined; reasoning by analogy, proof by dictum and special pleading are apparent at every turn, while control experiments and repeated observation never occur. But in this literature he does find much of interest; endless evidence of shrewd observation; a wealth of human study; and style of reporting which seems to combine literature, medicine and religion in an odd, persuasive way. He is told, however, that the only way really to know psychoanalysis is to be analyzed oneself—a matter of an hour a day for some 10 or 15 months—and that the fee will be from five to twenty dollars for each one of the 200 to 300 hours. In spite of the time and the money, enough experimentally trained psychologists have subjected themselves to analysis and reported the process so that we now have a working basis for applying the scientific method to the psychological aspects of psychoanalysis.

We now know that there are three major themes in psychoanalysis which must be carefully kept separate in our thinking. First, psychoanalysis is a special form of therapy, that is, a method of healing; second, psychoanalysis is a recondite, theoretical, philosophical system very similar to a religion; and third, psychoanalysis is a relatively definable, repeatable, partially controlled psychological experiment during which a scientific method of sorts is followed. I repeat, in order to think scientifically about psychoanalysis one must differentiate therapy, theory, and method.

Actually, all of psychoanalysis is oriented about therapy. This grows out of the historical fact that Freud's discovery which he called psychoanalysis was and is a psychotherapeutic procedure which produced beneficial results when applied to hysterical and neurotic patients. The chief medical interest in psychoanalysis remains in the therapeutic applications since therapy is the *raison d'etre* of the physician.

After the initial discovery of this therapeutic procedure, Freud's thinking followed two lines. The first was that of explaining the psychological phenomena revealed by the therapy. The second was the standardizing and perfecting of the technical procedure and methodology followed in the therapy.

Freud stated that he avoided reading widely concerning previous philosophical or psychological views related to the phenomena he discovered in analyses. This he did to avoid prejudice and preconceived ideas. To explain the phenomena of psychopathology he built up a system called "metapsychology" which is a philosophical speculation regarding the mind; a type of speculation which cannot be verified by direct experience. To most psychologists this metapsychology is part rational, part irrational, part dictum, part analogy and part based on clinical observation. It has all the strength and weakness of a new viewpoint essentially derived by one brilliant thinker who declined to contaminate his thinking by reference and conference with the work and thinking of his predecessors and contemporaries in the same field. (There was some reference to ethnology and folklore but almost all other contemporary science was neglected.)

Essentially it is this psychoanalytic metapsychology which has been and remains the center of controversy. Practically every scientist of every scientific discipline who has interested himself in psychoanalysis has found much that was stimulating and much to criticize. Most psychiatrists and some neurologists have completely accepted psychoanalysis including metapsychology, while most other medical specialists have either condemned or have reserved their judgment of the theory and practice of psychoanalysis.

Where the acceptance of the metapsychology has occurred, it has been due, I believe, to a felt need for some system of explanation which would provide guiding principles with which to deal with psychopathological phenomena. There does not exist any satisfactory or inclusive systematic frame of reference for explaining many psychopathic phenomena other than that provided by the metapsychology of Freud. Evidently, whether one wishes to acknowledge ignorance or to accept the highly speculative and esoteric system of Freudian metapsychology is a matter of personal preference plus the exigencies of the immediate problems.

Metapsychology may be dealt with by the speculative philosopher the theoretical psychologist and the theologian. The therapeutic aspect of psychoanalysis is a matter for the clinician to accept or reject.

The experimental procedure followed during a series of analytic hours does come into the compass of experimental psychology. It is worthwhile for the experimental psychologist to look into this aspect of psychoanalysis, using the usual criteria imposed by the scientific method in the hope that the differentiation between fact, theory and analogy may become more clear. We must inquire into the process of analysis to ascertain, where possible, how the psychological phenomena are related to the analytic process and method, disregarding all metapsychology and other theory; in brief, what is done, how it is done, and what happens.

At the very beginning of a psychoanalysis the analyst asks the patient to be absolutely frank and not to withhold intentionally anything that crowds into his mind and to try to overcome gradually all influences that may prevent his reporting certain of his thoughts or memories to the analyst. The patient must learn to talk without reticence or restraint, allowing his mind to proceed with its thoughts in a free fashion, reporting whatever comes up no matter what the nature of the thought may be. This is a process of uncontrolled or free association which must be learned.

During each analytic hour the patient reclines on a low couch while the analyst sits in a chair at the head of the couch, out of the range of vision of the patient. The physical surroundings, the hour of the day, and the patient-analyst relationship are kept constant. While the patient reports the analyst writes in a notebook. For the most part the analyst limits his verbal comments to the pointing out of relationships or to an occasional suggested interpretation. His usual reply to a direct question is "What do you think?" or "Why do you think so?" When advice is requested he deliberately restricts the scope of his answers. He points out to the patient the necessity for making no major decisions or changes in his routine of existence during the analysis. The set-up is one which emphasizes and tends to extend the psychoanalyst's prestige and power in the mind of the patient. Every effort is made to maintain the period of analytic sessions under as constant conditions as possible. It is true that individual practitioners modify this process to fit either their own ideas of the patient's needs, but the conditions for any one individual are relatively constant.

During 1940 the *Journal of Abnormal and Social Psychology* printed a symposium on psychoanalysis as experienced by 10 psychologists who have been analyzed. The motivation leading to entering an analysis, the type and length of analysis and the adequacy of the report of the psychologists varies widely, but enough similar material is reported so that we may, with some assurance, describe certain common psychological phenomena which occurred during these analyses.

An outstanding phenomenon occurring especially during the earlier months of analysis is *anxiety,* guilt feeling, or generalized apprehension. This is a pervasive unrest which invades all of the patient's thinking and interpersonal relationships. This anxiety may attach itself to some specific event or person, or it may be generalized without attachment to anything specific. In its unattached form it colors existence in the same way that colored glasses tint the visual field. When this anxiety does attach itself to some event or person, it does so in what seems to be an irrational fashion. Only after considerable free association does the attachment become plain and the connections obvious between the feeling-tone and the point of attachment. A further characteristic of the anxiety experience is the way in which it may enter into all varieties of judgment and conscious behavior in a dissociated fashion.

The mental state called *resistance* is manifested in either of two circumstances. Either the patient finds himself continuously blocked in his free associations and can think of nothing to talk about, or so many things enter his mind that he is unable to select which topic of conversation he shall follow. The first variety of resistance grows out of a necessary selection of topic of conversation. If one has told one's autobiography and then elaborated on it, one may find oneself at a loss for conversation. Talking about present-day events, business, science, etc., is not encouraged. The patient has a distinct feeling that he is wasting his time and money in such impersonal conversation. Yet the every-day conversation of most of us is non-personal. He therefore has to break down an habitual mode of social response and substitute for it a childish, irrational, emotional, personal monologue. This is difficult and annoying. In the process of learning this substitution the second alternative just mentioned may occur; namely, the difficulty in selecting the topic to be discussed. For example one may have had a dream which perseverates in one's thinking, one may have remembered some emotional event from one's childhood, and one may be thinking of some recent personal emotional experience. All three lines of thought seem to run along apparently independently and simultaneously. Which of the three should be reported? Anxiety plus resistance, that is trying to talk with nothing to say or too much to say, while paying for the privilege, constitutes a real emotional frustration.

Unconscious memories are said to be brought out by analysis. It is true that many forgotten events are recalled. However, to the professional psychologist this seems neither strange nor difficult to explain. If one revisits a scene once familiar but long forgotten, the associated detail of forgotten memory which comes into the foreground of consciousness is well known and recognized. If one sets oneself the problem of reconstructing the past in as great detail as possible it can be done with surprising completeness. When one strives to recount childhood memories and emotional events, the describing of certain memory traces leads on to many other details which were thought to have been forgotten. As a consequence one is able to recall many occurrences which somehow seem to be related to the formation of adult attitudes and habits. One is able to understand without too much difficulty why these forgotten events are influencial. There is no particular need to describe this phenomenon in terms of unconscious motivation or unconscious dynamics. It is a redintegrative process in which adult responses carry with them an emotional tone and character identical or very similar to that which occured in many childhood experiences. The concept of redintegration, that is, the arousal of a response by a fraction of the combined stimuli which originally aroused it, has a long history and quite sufficient experimental evidence to make it scientifically acceptable.

In this process of recalling forgotten memories one necessarily assumes a childish attitude and childish way of thinking. This has been termed *regression.* That it is true regression I doubt. It is a specialized attitude necessary for adequate recall.

A special feature of these forgotten memories involves the childhood emotional reactions toward one's parents. This goes by the analytic term of *Œdipus Complex* or Family Romance.

Since the child is so emotionally dependent on his parents and since it is necessary at the time of adolescence to attain the emotional independence of an adult, there can be no doubt, that there are many emotional conflicts and frustrations which are only partially resolved and which persist, coloring adult attitudes. These particular phenomena are not difficult to understand. It is merely that an emotional tie did exist between the child and his parents and that it was only partially dissolved.

This leads to the whole notion of the *dynamic unconscious* which is a basic concept in metapsychology. Again it seems that for the most part this is but the residuum of forgotten emotional experiences which were not resolved at the time of occurrence and which are, or may be, still active in the mental life of the individual through the process of redintegration.

Dreams and *dream symbolism* have started a great deal of controversy. For several reasons dreams are excellent taking-off places for free association. They are fantastic, childish, disconnected, and illogical. Their recounting sets up disconnected and illogical thinking. Since they are made up of a mixture of elements there are many lines of departure for free associations. The dream attitude of unconfined reverie is exactly that which is most conducive to free associations of a wandering sort which may bring out forgotten memories and associations which frequently furnish the clue leading to the explanation of a style of reaction. The interpretation of the dream by the analyst is partly systematic and partly trial and error. It does seem that many elements of thought carry with them common secondary meanings which somehow involve themselves in the dream. These secondary meanings provide a convenient basis for interpretation or pointing out of connections for further free association. Thus, the analyst may say, "It seems to me that your dream means such and such." If this notion is acceptable, the patient can proceed with his associations and perhaps gain further insight concerning some line of thought. If the suggestion is unacceptable, the patient can reject it and there is nothing lost.

Transfer is said to be an invariable mark of a true therapeutic analysis. Psychologically transfer and conversion, as described by James in his book, *The Varieties of Religious Experience,* are approximately identical. In essence, transfer is a psychological unloading of the patient's troubles on the analyst, accompanied by an emotional over-evaluation of the analyst. He endows the analyst with extraordinary or supernatural wisdom and hopes that the analyst will be able to point out quickly the solution to his troublesome thoughts and habits. In a sense, transfer is a falling in love with the analyst. The erotic nature of this transfer is mentioned with sufficient frequency to convince one that it is an actual occurrence in many cases. In other individuals the transfer is the assuming of a child-parent relationship by the patient. The patient is putting himself into a childish attitude, trying to bring back forgotten memories. Still another aspect of transfer is the assumption of the same relationship which exists between the priest and his parishioner. The priest is, by common consent, endowed with certain supernatural qualities and powers. The patient assumes the role of the parishioner and trusts that the analyst will lead him to reward and mental tranquility just as a priest might do.

Many reporters have commented on the character of *insight* gained during their analysis. The patient feels that he understands himself better, is better able to cope with his problems and is less disturbed by troublesome social and personal situations. This is usually attributed to the acquisition of insight. There is nothing mysterious about this. If one has recounted one's autobiography to some understanding listener and has had comments made on the autobiography and associations pointed out which were not previously obvious to him, he is certain to gain some insight. Insight, then, means an educative process in which many unverbalized attitudes and effects become verbalized and satisfactorily integrated into the entire personality.

This entire process of enforced free association, under the circumstances which I have described, leads to anxiety, irritability, resistance, regression, transfer, etc. All these are symptoms of *neurosis*. If one defines neurosis as the persistence of non-adjustive emotional attitudes, constantly recurring and leading to inadequate behavior, then the analytic process must by its very method create a neurosis and insofar as the individual resolves his emotional inadequacies and gains insight, he will be freed from the neurosis.

All these phenomena occur with a great deal of regularity during analysis. There are other phenomena which occur in lesser degree or which are prominent in one case but not in another. To the best of my belief, there is no need to have recourse to metapsychology to explain these additional phenomena any more than there was to explain the phenomena which I have just discussed.

The peculiar mental experiences which I have mentioned were commented on by most of the analyzed psychologists who reported their experiences in the symposium. Most of the reporters saw no reason to believe that their experiences transcended the explanatory principles sufficient in general experimental psychology. Several reporters mention how the psychoanalytic procedure itself served as the basis of certain of their experiences. It is my belief that most of the phenomena brought out by psychoanalytic procedure are the combined result of the method plus the personality under study. This is a combination in the same sense that the behavior and experience of a man who is alcoholically intoxicated is a combination of effect of alcohol and his individual personality. Hence, it seems to me that most of metapsychology is an unnecessary speculative construction. The dynamic unconscious is more easily understood in terms of childhood memories of which the effect and the affect went unrecognized until more closely surveyed. The Œdipus Complex turns out to be a question of family ties. Transfer is a combination of prestige suggestion and adoption of a childish attitude. Dream work and dream symbolism become a form of waking acceptance of free associations which may or may not be reasonable. I am, in effect, saying that there is nothing mystical or esoteric brought out by analysis. The use of ordinary scientific method would clarify those points which are still obscure and general experimental psychological principles will furnish all of the necessary explanations without recourse to metapsychology.

There remains the question of the claim that sex is the basic, driving, motivating urge of human life and experience; that from the libido (undifferentiated sex energy) comes all of the basic developmental forces of personality growth and function. Again the evidence shows that the metapsychology does violence to fact. Freud's original therapeutic measures were developed through the study of hysterical and neurotic patients. Other psychopathologists had shown before and later investigators have shown since that the leading psychological defect in these patients is an emotional inadequacy marked by a generalized fear called anxiety or guilt feeling. Free association starting from the anxiety symptoms frequently leads to forgotten sex memories. Freud then held that this sex element came before and was more important than fear. It seems to me that when one is dealing with the memories of the undifferentiated emotions of infancy or early childhood, there is little reason to name the experience sex or fear or anger. It was, and still is, more socially acceptable and more scientifically "proper" to discuss fear than to discuss sex. To cut through a long, involved, and often fought-over hypothesis, I believe that one can safely say that out of an undifferentiated emotional reaction mass of infancy develop all varieties of emotional reactions, affects, sentiments, etc., of adult life, and that one important differentiated system among several others is that which we call sex. The overwhelming emphasis on sex in Freudian metapsychology does violence to evidence admissable under the rules of scientific method.

Next let me mention briefly the impact of psychoanalysis upon, and its importance to, social science and anthropology. Let no one doubt that psychoanalytic concepts are used more in anthropology, sociology and economics than are those of experimental psychology. The academic psychologist just never got around to providing experimental evidence or even adequate philosophical speculation concerning the development of the personality, character and temperament of the entire man. In similar fashion, nothing was prepared for the physician, psychiatrist or neurologist when he came to deal with the problem of the entire organism. Psychoanalysis did present, through metapsychology, a workable hypothesis. The scientist working in any field must accept and use the tools and constructs of those working in the scientific disciplines basic to his own field. For the most part, the psychologist must accept the findings of the physiologist, biochemist, and endocrinologist as factual and well established. The fact that psychoanalysis was and is a mixture of fact and unsupported theory is unfortunate for all concerned, but most of all it is unfortunate for the psychologist, since he should have provided and ultimately he must provide the claried and unified scientific basis of explanation in this realm of knowledge.

Meanwhile, our colleagues in the social sciences and in anthropology accepted psychoanalysis as a factual system and are attempting to orient their scientific efforts in these terms. This has led to endless confusion. It is as if the physiologist asked the chemist for a method of study and an hypothesis to explain the chemical action of the pituitary gland and was told that no exact chemical knowledge or methods of study existed but that a Russian pathologist had made some interesting observations on pituitary function and developed an esoteric chemical theory which no chemist had verified. Should the physiologist wait until chemistry does develop a method or explanation, or should he accept the brilliant speculations of the pathologist?

In conclusion, allow me to summarize the factual contributions of psychoanalysis to modern psychology. In the first place, psychoanalytic investigation forced psychologists to recognize and to inquire extensively into the role of infantile and early childhood thought and fantasy on the mental life of the adult. Here the contribution was direct from psychoanalysis to psychology and of first importance. This constitutes a rapidly developing body of knowledge. Closely associated with this is the second contribution which the psychoanalysts have called the dynamic role of the unconscious. That forgotten memories, emotional experiences put out of mind, childhood fantasies, the prohibitions of childhood training, etc., etc., do influence thought and action and in ways not readily recognized by the individual himself, is now apparent and psychologically accepted. This field is being studied by the more usual psychological methods and in due course the mysterious dynamic unconscious will probably turn out to be but part of a section of the learning process. A third contribution has been the forcing of the attention of psychologists to the psychology of sex and sexuality. This may not be comfortable but as psychologists we must deal with the biological organism as it is and in all its aspects. We can no longer continue the cultivation of polite blind spots protected by the dictum that such topics are not the proper field of psychological study since Wundt did not include them in his experimental program announced in 1880. Lastly, the psychoanalysts, by setting up a theory and explanation of the structure of personality which was and is unacceptable to most psychologists compelled psychologists to initiate the scientific study of personality.

PSYCHOLOGY AND DEFENSE

ROBERT M. YERKES

Professor of Psychobiology, Yale University

(Read April 25, 1941, in Symposium on Recent Advances in Psychology)

THE caprice of circumstance, not personal preference, dictated my subject in this program on "Recent Advances in Psychology." Ordinarily I should have chosen to sketch the history of Comparative Psychobiology in the current century. Instead, as one-time commanding officer of the first company of military psychologists in the world, I speak of relations of my science to the art of defense. By psychology I shall mean the science of behavior-experience; by defense, protective, preservative, and constructive activity. Using military psychology and its notable exhibits as my point of departure, I shall later abandon the rôle of historian for that of social planner and discuss relations of psychology to defense against such internal weaknesses as ignorance and superstition, selfishness and cowardice. For what we are individually is more important than our immediate objectives, since strength comes only transiently from needs induced by aggression, while it comes lastingly from faith and deep conviction. Looking through and beyond this world clash of cultures, I purpose to exhibit psychology as a basic science and profession, the applications of which are as varied as the problems of behavior and the potential significance of which for our civilization matches that of the physical sciences and their respective branches of engineering.

Ours is the tradition that life is sacred and that nothing human may safely be tampered with. Therefore, our preoccupation with material things, to the neglect of efforts to control human form and functions. It is not surprising, in view of this ancient bias, that mental engineering should have remained practically unknown to military men until our time. The discovery dates from our declaration of war in April, 1917. Promptly American psychologists decided to try to increase military effectiveness by bringing their information, techniques, and insights to bear on problems of manpower. As personnel specialists and psychological examiners, they assisted with the appraisal of recruits by supplying descriptive measurements of mental alertness, adaptability, and trade or other skills. Thus they contributed to the data for intelligent military selection (acceptance or rejection), classification, assignment, and training.

Aside from the solution of varied psychological problems for the army and navy, two principal lines of service were developed. The Division of Psychology of the Medical Department of the Army examined, rated, and reported on the mental caliber of nearly two million men, while the Committee on Classification of Personnel of the Adjutant General's Department trade-tested, rated, and aided with the assignment of other hundreds of thousands.

As I have elsewhere expressed it:[1]

> The theory of psychological service was that human factors should be appreciated, measured, and intelligently used, that so far as feasible chance, personal whim or bias, and convention should be replaced by action in the light of reasonably accurate and thorough information. In a word, that the army should utilize what may be called "human engineering," just as it attempts to utilize other forms of engineering which have to do primarily with non-living things.

These novel services, of which the development inevitably aroused both strenuous opposition and enthusiastic support, necessitated the preparation of methods and equipment, planning for organization and integration of activities, and the recruitment of hundreds of young men qualified for training as military psychologists. When the Armistice became effective in November, 1918, between two and three hundred psychological specialists were serving in our military establishments. As a result of their labors, psychotechnology had demonstrated values and potentialities of development which commanded the lively interest and respect of military leaders.

The World War saw no developments of military psychology or other personnel services in any other country which were remotely comparable with what occurred in the United States of America. Germany, late in the war, tested certain psychological procedures for personnel selection, with favorable result; Canada concentrated on improvement of methods of military training and reëducation; and England utilized her psychologists chiefly to supplement the labors of medical officers. Ours, because of superior professional resources and less acute risks from aggression, was an altogether unique opportunity for pioneering, the prompt improvement of which yielded signal successes.

---

[1] Yerkes, Robert M. "The Rôle of Psychology in the War." Chapters XX–XXI in *New World of Science*, p. 358, 1920 (New York, Century Co.). [Reprinted as separate, 1941.]

The war's end in 1918 signaled the return of military psychologists to their customary occupations. Our army and navy retained records and memories of their services, but not even a cadre of specialists. Following the Armistice it was wisely decided, in the light of previous experience, to center all psychological activities hereafter in the Department of the Adjutant General instead of permitting them to continue divided. But unfortunately provisions were not made for the continued development of military psychology as mental engineering within the army. This chapter of service ended abruptly. There were universal sighs of relief, and the Congress promptly acted on the assumption that the war to end wars had actually accomplished that purpose. Isolation had its field day, and instead of capitalizing our precious experience and consolidating the encouraging progress which had been made in mental engineering, we virtually ignored the opportunity in favor of our commercial pursuits. For twenty years military psychology remained for us merely a matter of historical and academic interest.

Not so in Germany. Awakened to the practical values of psychology by a few demonstrations at home and the startling achievements in America, the leaders of the Reich promptly included psychologists in the organization which was designed to reëstablish Germany's might.

To understand the rôle of psychology in postwar German developments, it is necessary to see the situation clearly and in its entirety. The nation had suffered a technical defeat at arms while remaining unconquered internally. Following the unaccepted defeat came a peace unwise in its terms, and even more so in its enforcement. It must be reckoned one of the tragic failures of civilization, since an unprecedented opportunity to safeguard the world against future wars was sacrificed to narrow-minded selfishness, fear, and shortsightedness. Instead of being crushed by the severity of the Treaty's terms, Germany was goaded to determination and superhuman effort to recover her strength and turn the tables.

The dominant German thought-pattern, fostered by the Treaty of Versailles, includes a philosophy of war in which scientific and technological discoveries hold a conspicuous place. Warlikeness is considered admirable, something which is characteristically German, to be prized, cultivated, used for national aggrandizement.

It was inevitable that the idea of mental versus material warfare should seek and find place in this philosophy of war. First proposed as hypothesis that feelings, thoughts, and attitudes might be used as instruments of warfare, the idea soon was put to the test in the civil contest between democratic and Nazi ideologies. It is said that the Nazi leaders thus proved to their satisfaction the effectiveness of mental weapons.

From the acceptance of mental warfare as real to the development of the concept of total war was an easy step. Total war may be defined as the employment of all available material, mental, and spiritual resources to achieve national objectives. It is unannounced, timeless, spaceless. In warfare as thus conceived, psychology and its applications become as indispensable as are the physical sciences and their technologies. We must now inquire: How has psychology functioned in the rehabilitation of Germany and in the emergence of Nazism as a threat to other cultures and ideologies?

The German Army of one hundred thousand men, which was authorized by treaty, was made up of volunteers who enlisted for twelve years. For their selection, available psychotechnological methods of appraising military qualifications were employed. When in 1935 opportunity for further militarization appeared, this small army of carefully picked and thoroughly trained men provided the organizers and leaders for a sudden and numerically vast expansion. This extraordinary achievement was possible because unusual military experience and acumen were reinforced by the skillful scientific handling of problems of manpower as well as machine-power. Military psychology thus gained full recognition in Germany and came to command further opportunities for the demonstration and development of values. Henceforth, human engineering took its place in the Third Reich as complement of physical engineering.

For a decade at least, Nazi Germany has been molding psychology as an instrument to further social development and enhance prestige, influence, and military power. Psychologists have been made specialist servants of the state, and remilitarization stands as only one of many phases of Nazi evolution in which they have proved their usefulness. Characteristically thoroughgoing study of the problems and risks involved and patient planning prepared the way for the present system of state psychological service. Already hundreds of young men have been trained for the various branches of this highly technical service. Selection for such opportunities is made on the basis of personal characteristics and professional promise, and the course of training extends over years and is comparable with that required for other learned callings. Back of a personnel which is impressive in size and effectiveness is a highly equipped complicated organization.

The Central Psychological Institute is in Berlin. It combines the functions of research laboratory and professional school. Psychological problems of propaganda, morale, personality analysis and appraisal, educational training, industrial efficiency, fatigue, and similar concerns of human engineering, occupy the staff. Military psychology is one only of the divisions of the subject in which a student may choose to specialize. Those who elect this specialty have been required heretofore to work two to four years as probationers in central and field laboratories or stations for military psychology and in army training camps. Presumably the procedure is very different under the pressures of active warfare.

The objectives of the Central Institute are furthered also by university laboratories and branch establishments in army corps. In all, research is the order of the day. Methods for the observation, measurement, and control of human behavior are continuously being devised, adapted, developed, tested, and evaluated. Pressure and incentives to progress through discovery are not lacking. It is said that the Nazi army psychologists have come to rely on methods of personality analysis and the judgment of examiners instead of on the results of intelligence tests; that characterology, temporarily at least, is in the ascendant and psychometry in eclipse. The truth, however, appears to be that intelligence tests have been supplemented instead of abandoned, and that a large and steadily increasing number and variety of observational techniques serve for the appraisal of the individual.

The degree of thoroughness and comprehensiveness of psychological examinations for important posts or occupations is indicated by the time devoted to them. Thus the examination of a candidate for commissioned rank in the army requires two days, with an intervening rest day. It is conducted by a board made up of civil and military psychologists and army officers. The ordeal includes several kinds of tests for intelligence, mechanical aptitude, will power, emotional stability and control, leadership. The examiners do their best individually and collectively to discover and rate important mental traits and to obtain a practically serviceable picture of the personality of the candidate. American military psychologists of World War experience will marvel at the contrast, for our officers of staff and line considered it generous to allow minutes instead of hours or days for the psychological examination of a recruit or a candidate for training as an officer.[1]

In the existing emergency, whether or not technically at war, we assuredly are not at peace, even in our own minds. Moreover, we are no longer as in 1917 pacemakers in military psychology. Germany long ago assumed the leadership, and at the moment we are but halting followers. As in armament production and the making of soldiers, so in psychological defense, we plan, and are under compulsion, to accomplish in a few months what the Nazis have labored over for a decade. It is to be hoped that we shall honestly and resolutely face the facts, instead of making believe that we are in all essentials superior to the totalitarian peoples and their means of aggression and defense.

The hope and expectation that we can greatly exceed Nazi speed of militarization and psychological preparation for total conflict are neither idle nor presumptuous. For although among us the development of military psychology virtually ended with World War I, the progress of the science of psychology and of many of its technical applications since 1918 has been phenomenal. Our available resources of personnel, centers for research and professional training, equipment and useful techniques greatly exceed those of any other nation. It would seem that in the existing stress and strain we should by all means and immediately conscript our invaluable psychological assets, enter without prejudice or qualms into the fruitful psychotechnological experience of the Nazis, taking over whatever they have found good and improving on it, and above all and continuously, concentrate our ingenuity on discovery and invention. Thus with sufficient determination and good fortune, we might succeed in compressing the worthwhile progress of a Nazi decade into a year of American effort.

---

[1] For sources of information concerning German Military Psychology, see "German Psychological Warfare: A Critical, Annotated, and Comprehensive Survey and Bibliography," issued by the Committee for National Morale, New York City, 1941.

The current picture may be sketched boldly. Psychologically we are ill prepared for participation in world conflict. Much excellent work in mental engineering is in progress in our academic institutions, industrial laboratories, under the auspices of the National Research Council, the United States Employment Service, Civil Aeronautics Authority, and such governmental departments as Agriculture, Army, and Navy. The conspicuous shortcoming is the lack of a comprehensive plan of organization for the proper relating of interests and the achievement of common objectives. Work is relatively uncoördinated, and individualism seems to be outdoing itself and in many directions needlessly delaying progress. Thus far as a people we have refused to face honestly the urgent necessity for selection of personnel on objective grounds of ability and qualification. We give support grudgingly to the principles and practices of civil service and to our varied forms of merit system. This popular attitude markedly affects the practices and progress of our military establishments. Especially prejudicial to the effective utilization of manpower in our present urgent need is the failure, to the date of writing, of the army and navy to provide training schools for military psychologists and other personnel specialists. For what will take longest and are of incomparable importance so far as efficiency of psychological service is concerned are the wise selection and thoroughgoing training of an adequate supply of specialists in military mental engineering. These deficiencies or temporary neglects are grounds of criticism but not excuse for pessimism, for if we choose we can quickly transform a chaotic situation into one of strength and symmetry.

With this general review of the rôle of psychology in World

Wars I and II, I shall pass to discussion of the relations of the science and its technologies to other than military modes of defense of human life and its quality. The theme will be: Psychology as profession. We of the New World persist in thinking of warfare as an outmoded method of resolving disagreements, which will inevitably be replaced in the course of the evolution of civilization and with progress in the socialization of nations. It is because of this assumption and faith that I can devote the major portion of this address to the consideration of emerging social needs, the satisfaction of which depends chiefly on our understanding of human behavior and experience and our ability and willingness to control them. Doubtless at this juncture we should concern ourselves rather with the improvement of man and his lot than with attempts to preserve any particular national or racial group or culture. For a world shared peaceably and coöperatively by several nations or ethnic groups with a common culture is not unimaginable, although it now seems a remote possibility.

There is abundant evidence that social adjustment has failed to keep pace with material progress, and that mechanical invention, by transforming our space-time relations, has brought about acute social maladjustments and thus created needs for professional services which are not available. Even those who have neither the will nor time to study the situation realize its gravity and admit that something radical must be done if social progress is to be continued without worldwide destructive revolution.

In our gropings for assistance we have looked naturally enough to the so-called learned professions. Probably the vast majority of us still think in terms of three great professions: religion, law, medicine. Yet the notion is archaic, for actually thirty probably would come nearer the mark than three, and in every one of the professional callings, as contrasted with trades and other occupations which depend primarily upon mechanical skills, there are notably learned persons. It is in fact difficult to make useful distinction nowadays among the terms "vocation," "occupation," and "profession." The latter is commonly accepted as the least inclusive of all and especially applicable to those callings which require book-learning, if also apprenticeship and the acquisition of skills. Measurement quite likely would reveal that teachers in institutions of higher learning now rank in learning well above clergymen, lawyers, and physicians. Practically, which occupation ranks first matters little, but it is of great importance that the increasing extent of the diffusion of knowledge and the social significance of this phenomenon be recognized, and that we steadily increase the effectiveness of the utilization of this incomparably valuable resource.

The range of professional functions and of social needs extends from the reparative or restorative, typical of medical services, to the constructive and creative, as exemplified by the divisions of physical engineering. Between the human needs which are partially met by the physician and those which the clergyman or priest is expected to satisfy, there appears a great gap. Its importance has been accentuated in recent decades by the modifications of life necessitated by rapid change in environmental relations. Old needs have been intensified, new ones created, and a large part of them are unmet by existing professional developments. In many instances these needs are ill-defined, and our attempts to satisfy them are little more than gropings among our intellectual furnishings. It is entirely clear that the old professions must be adaptively modified to relate them more nearly to present-day needs and that new professions, perhaps in considerable number and variety, must be recognized, or as necessary created, developed, and safeguarded by educational requirements and appropriate legal certification of their practitioners. Thus only can they be given suitable social dignity and prestige.

Under the pressure of demand for services which lack a recognized professional specialty, both physicians and clergymen have tended in recent decades to extend their human ministrations into areas which they are not adequately prepared to enter. Medicine, for example, has reached into normal life and advised on problems of behavioral development and social adjustment which are primarily psychological or educational, and religion has extended from the realm of the normal into that of the abnormal. Illustrative instances are the adventures of psychiatrists in the application of psychoanalysis to individuals who have no more serious ill than the desire to be analyzed and the means to pay for the service; and those of clergymen who officiate in church clinics to which come, in addition to normal persons who happen to have spiritual difficulties, many with bodily and mental disturbances and diseases. In the one instance, medicine is encroaching dangerously on clinical psychology, and in the other, religion is venturing into the sphere of the physician or of the clinical psychologist. It is reasonable to maintain that unless medicine is prepared to include psychology among its basic sciences and religion to add several sciences to its curriculum, such extensions of sphere of service are likely to be unprofitable to the individual in need of dependable advice and to society. Professional competence of necessity implies knowledge, special training, and experience to enable the practitioner to deal skillfully and wisely with a restricted range of human problems. The physician is held responsible for mistakes or malpractice due to ignorance, carelessness, or dishonesty, and obviously the lawyer, clergyman, educator, or other professional specialist should similarly be held accountable to society.

The service gap which has been noted between medicine and religion relates primarily to psychology as a scientific discipline because the conspicuous needs which it represents have to do with the facilitation and intelligent direction of the development and current behavior-experience of the normal person. In this connection it would seem that psychology must stand as a basic science for such universally desirable expert services as the guidance and safeguarding of an individual's growth and development, education and occupational choice, social adjustments, achievement and maintenance of balance, poise, and effectiveness, contentment, happiness, and usefulness. There is indeed a large uncultivated professional area which should speedily be occupied by the psychotechnologies and related forms and aspects of human engineering.

If a digression may be pardoned, I should like to consider somewhat further the matter of definitions. It happens that "psychology" has proved a troublesome word; in many quarters it has caused considerable annoyance. It is not easy to understand the prejudices against the term, nor to see why it should be abandoned for other modes of expression. So considerable was the opposition to the word during the World War that in the United States Army effort was made to avoid its use by substituting the word "personnel," and for the technical term "intelligence," the word "alertness." Similar experience in Germany during the present conflict is said to have motivated search for substitute words or phrases. It is undeniable that the terms "psychology," "psychiatry," and "psychoanalysis" are frequently confused, misused, or even assumed to be synonymous by well-educated, intelligent persons. On the lower level of intelligence and information, psychology is sometimes confused with such special and diverse interests as psychic research, telepathy, spiritualism, Christian Science. Although to some it may appear an insult to intelligence and literacy, I venture to present the generally accepted definitions and distinctions which apply to psychology, psychiatry, and psychoanalysis.

Psychology is a basic natural science which is concerned with mental processes and their behavioral expressions. Previously I have briefly defined it as the science of behavior-experience. In my opinion it is one of the most important general divisions of biology. Psychiatry is not a basic science, but instead a medical specialty which is concerned with problems of mental defect, disturbance, and disease. Its most closely related areas of science are known as neuropathology and psychopathology. In view of these widely accepted and very definite distinctions, it seems grotesque that the terms "psychology" and "psychiatry" should ever be confused.

As for psychoanalysis, it is neither a division of biological science nor a recognized medical specialty, but instead, at one time, a method of discovering and observing behavior-experiences which are not readily accessible under ordinary conditions of self-observation, and, again, a form of therapy. As a scientific procedure it may be used by the psychologist or physician as a research tool, and by the psychiatrist either as a way of studying the nature of a patient's behavior-experience or of treating the individual. In popular usage, the word "psychoanalysis" unfortunately carries a burden of Freudian assumptions and interpretations which logically belong in psychology or psychopathology. As a matter of fact, a new word is needed to designate Freudian and related doctrines and systematizations of fact in distinction from the method which is appropriately named psychoanalysis. It would indeed be a blessing without disguise if what is valuable in these psychoanalytic accumulations might be assimilated by psychopathology and psychiatry, leaving psychoanalytic method to be used, like other legitimate introspective procedures, for scientific or therapeutic purposes, but always and only by trustworthy and competent persons.

The existing learned professions meet but a fraction of the human needs of which we are now aware. From the sciences of physics and chemistry there have developed branches of engineering which enable us to control and use effectively many features of our physical environment. From certain of the biological sciences have sprung such services of healing, repair, and protection as are made possible by the professions of medicine, surgery, hygiene, sanitary engineering. But from psychobiology, psychology, and the social sciences, corresponding developments within the sphere of human engineering, and planned to meet the needs of the normal healthy person, exist rather in promise or incipiency than in fulfillment.

It should aid us in visualizing the possibilities of expert assistance in living normally to list a few of the emerging psychological and social services which in the future may rank as professional callings. They represent a group or constellation of service developments for which we have no other term than psychotechnologies. In the future the enlightened individual, with lively self-interest and social-mindedness, may desire the availability of such special guidance in normal development and adjustment. He may even employ a consultant and adviser in psychological and social matters who in deference and esteem may take his place beside the family physician.

This is what one might wish to happen by way of guidance in self-realization and social usefulness. Thoroughgoing examination within the first three years of life by a psychologically trained specialist in child development, followed by appropriate report and advice concerning care and training. Comprehensive psychobiological examination and individual appraisal at about ten years of age by an experienced clinical and educational psychologist whose report and prognosis relative to developmental potentialities and educational needs might be used as basis for tentative judgment as to whether the individual should expect to make his living by brawn or brain-power.

Assuming capacity for highschool training, there should be available in the last year the services of an educational counselor, whose objective study should supply data for final recommendation and decision between leaving school for a trade or other occupation or continuing educational training toward some learned profession. In the latter case, advice should cover also vocational possibilities and a desirable college course.

Sometime during the last two years of work in college, the subject should be carefully studied by a vocational counselor, who in the light of previously accumulated information and his immediate findings should offer further vocational prognosis and safeguard final choice of vocation, since this might prove the most important decision of the individual's life. If thereafter some professional school, such as medicine, engineering, psychology, education, were entered, there should be available an experienced educational psychologist to advise concerning and direct the course of training and the process of self-development.

Still farther along, when marriage became a question and possibility, a specialist in marriage consulting should be sought, with confidence that he would bring to bear on the problems at issue knowledge of the characteristics of the prospective mates and pertinent facts of eugenics, genetics, social psychology, and sociology, and from such information and his wisdom offer prognosis and advice which would enable the individuals to reach an intelligent, socially minded decision.

It has of course been assumed in connection with this list of desirable, although still novel, professional services that it was the good fortune of the subject not to require the services of a child psychiatrist because of congenital or accidental mental defects or disturbances, or later in life of a neuropsychiatrist or a clinical psychologist to deal with mental disease or social maladjustment. This prayer offered as an individual expression of desire for the availability of professional services over and above those which customarily are provided by physician, lawyer, or clergyman rests also on the assumption that the organic stuff of the individual promised reasonably normal and worthwhile existence. If instead, in the judgment of those professionally competent persons with whom the social responsibility of decision and action should rest, his being held no promise of complete, serviceable biological development, self-supporting independence, and usefulness, it would seem socially defensible that his life should be ended gently, without vain regrets or social waste.

If called before a court of my peers to defend the social practice here suggested, I should say simply that we of this country and culture have destroyed or impaired the effectiveness of most of the naturally operative factors and agencies for the elimination of the biologically unfit; that we have lacked the courage, wisdom, and resourcefulness to devise and put into effective operation socially controlled substitute procedures to safeguard human quality; and that in consequence of these social neglects we are steadily becoming less fit for survival and, incidentally, through the institutionalizing of those who are physically or mentally defective, insane, criminalistic, or for other reasons wholly incompetent and dependent, we have accumulated a degree of social responsibility and an economic burden which drastically limit opportunity for constructive endeavor. It is reasonable to maintain that our resources might be much more wisely used in the enhancement of normal life than in protecting and prolonging individual helplessness and misery. So far as I have been able to discover, our unwisdom in this matter is attributable entirely to sentimentality and a false conception of humaneness.

Granted the need of new services in the nature of human engineering to improve life, and of a profession which shall be to the normal man what medicine is to the abnormal or otherwise afflicted, what should we do about it? The proposal which follows is the chief excuse for this address. It is, briefly, that psychology be recognized as science and profession basic to the development of all branches of mental engineering; that professional schools of psychology, after the model of the best in medicine and physical engineering, be established; that the standards of admission be high and selection on the basis of ability rigorous; that four years' intensive work be required, of which those who hope to become practicing psychotechnologists shall devote two to basic science studies and the remainder to technological training which would correspond to the clinical years of a medical course; that graduates of such schools be awarded the professionally distinguishing degree of Doctor of Psychology, or some more appropriate designation; and, finally, that any graduates who may wish to practice a psychological specialty be permitted to do so only when licensed under appropriate federal and state laws.

In making this novel proposal for the constructive as contrasted with the military defense of man and his cultural achievements and strivings, I am looking beyond the present world conflagration to a period of reconstruction during which innovations are likely to be the unescapable order of the day and the fashioning of a new civilization a necessity. To say that I am speaking inopportunely or hopelessly in advance of the development of psychology as science and of popular realization of needs and opportunities for service may be the truth, but I offer in reply as battle cry of progress the old slogan, "Nothing ventured, nothing gained." For it is my conviction that unless we can promptly make up our minds to risk all in determined efforts to reshape our social, economic, psychologic world order, we possibly shall lose all. I have endeavored to make a recommendation, which, if wisely instrumented, might immeasurably increase the values of individual life, foster social-mindedness, and improve and stabilize social organization.

In the picture which has been drawn, psychology, as a major division of human biology, with varied potentialities of application to human life, stands as complement beside the great profession of medicine. Each from the point of view which I have taken should meet a wide range of urgent human needs; each should provide with increasing adequacy for several special forms of human engineering. Medicine's preoccupation has been, and might well continue to be, with the abnormal and pathological manifestations or vicissitudes of life and with scientifically based and intelligently directed efforts to make and keep the individual whole and sound; whereas psychology naturally would and should be occupied with the normal or typical personality, with questions of individual development, educational or other forms of training, vocational choice and activities, social behavior and adjustments, personal competence, contentment, and happiness. The potential values of increasingly intelligent attention to problems of normality, effectiveness, and fitness, versus those of defect, disease, and unfitness, must be obvious to every intelligent, open-minded person.

We of this civilization have reached a stage in structural and cultural evolution which renders it unnecessary that attention and effort be concentrated on negative measures of defense in order that the species may be preserved and perpetuated. Having escaped this primitive necessity, the opportunity is ours to think constructively and to further self-realization, self-development, and social endeavor by providing to the utmost for increase of knowledge and enlightenment and the development and utilization of professional services. Psychology, I have maintained, should be concerned as a profession primarily with the improvement of human nature and the enhancement of individual and social human values. That it is crude or in its infancy has no bearing on its ultimate and potential usefulness or the possibility and desirability of developing it rapidly both as fundamental science and as technology. It is my prophecy that eventually the general area of knowledge which has become the main topic of this address—and it matters little whether we designate it by the word ''psychology'' or otherwise—will take its place as one of the most important fields of scientific inquiry and learned professions.

MOTIVATION, LEARNING AND ADJUSTMENT

EDWARD CHACE TOLMAN

Professor of Psychology, University of California

(Penrose Lecture, April 25, 1941, in Symposium on Recent Advances in Psychology)

As you all know, Jung divides us into the extraverts and the introverts or, to use a pair of terms suggested earlier by Lowes Dickinson, into the redbloods and the mollycoddles. An extravert, or redblood, is one who orients outwards. He is directly responsive to external stimuli. He sees the world as it is. An introvert or mollycoddle, on the other hand, is one who orients inwards. He is responsive to external stimuli only in so far as they fit into his own preoccupations. He sees the world colored—and by his own personal biases. And I am an introvert. Hence when the invitation to give this lecture suggested that I present a summary of some recent movement in psychology, I was troubled. For only an extravert or redblood could do that. Only an extravert or redblood could honestly and objectively summarize any of the current advances in psychology—even in rat psychology. It seemed to me, therefore, that either I must refuse the invitation altogether or else I must ignore my conscience and, while pretending to give an objective unbiased account of some recent concepts, really utilize this occasion to put across some of my own distorted evaluations of those concepts. It is this latter course which I decided upon.

The words "motivation," "learning" and "adjustment" were chosen for the title because it was hoped that their general connotations are such as to suggest, even to the non-specialist in psychology, the general field about which I want to talk. It is the field which concerns itself with the problems of individual adjustment—the field with which the psychoanalysists and the other clinical psychologists have all been concerned. It asks the question: "Why on a particular occasion a given individual wants what he wants?" What are the laws which determine his wants; and how can these wants be made more appropriate than

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 4, JUNE, 1941

they so often tend to be? In attempting the answers, we shall find it helpful to consider successively four sets of concepts, namely: (1) drives; (2) values and valences; (3) beliefs, and (4) frustration mechanisms.

1. DRIVES

Man is to be credited, I believe, with a relatively long list of innate drives. These are to be subdivided, first, into: (*A*) the *biological drives;* and (*B*) the *social drives.* And the former are then to be further subdivided into: (i) the *appetites* and (ii) the *aversions.*

(*A*) *The Biological Drives*

(i) Consider, first, the *appetites.* Man has quite a lot of them. I suggest the following:

The Appetites

A maternal (or suckling of the young) drive
A nest-building drive
Thirst
Hunger
Sex
A general activity drive
An exploratory drive
A rest or sleep drive
A urination and defecation (in specific
  type of locale) drive
A play drive
An æsthetic drive

Practically all of these can be observed in chimpanzees as well as in men. And many of them have in fact been located and studied, to advantage, in even so lowly an organism as the white rat. The identifying feature of each of these appetites seems to be a specific and characteristic *consummatory response* —*e.g.*, suckling, hiding in nest, drinking, copulating, exercising, exploring, lying down and sleeping, urinating or defecating in a given spot, playing, engaging in sensory activity for its own sake. Such a consummatory response ends ultimately in a final state of internal satiation, whereupon the appetite subsides.

Further, the strength of any such appetite, at any given time, will be measured either in terms of the intensity of the consummatory response itself or in terms of that of a "preparatory" response which tends to get the individual on towards the presence of the necessary consummatory object.

Let me present briefly some of the experiments on white rats.

The Maternal Drive.—A recently parturient female when removed from her nest will (as demonstrated by Warden and his co-workers at Columbia) cross an electric grill many times in succession to get back to her nest and young. And the number of such crossings which she will make in twenty minutes can be taken as a direct measure of the strength of her maternal drive under the given conditions. And, again, Wiesner and Sheard in Edinburgh have shown that, if the young be removed from the nest and dropped into the cage at some distance away, the mother will be ready to retrieve them and carry them back to the nest many times over. And it was found that the number of such retrievings which she could be induced to carry out, could also be taken as a direct measure of the strength of the maternal drive at the given time.

The *nest-building* drive. A just pre-parturient or post-parturient female (or any rat—male or female—kept at a low temperature) will build a nest from paper strips, if given an opportunity. When the nest is done the animal hides in it. And Kinder, working in Stone's laboratory at Stanford, has shown that the number of such strips which the animal will use for building its nest will vary with the temperature, the nearness of delivery of young, and the like.

Thirst, hunger and *sex,* like the maternal drive, have been measured mostly by numbers of crossings in twenty minutes of the electric grill which will be undertaken by the animal to reach the appropriate goal-object.

The *general activity* drive has been measured by an activity wheel, similar to that attached to a commercial squirrel cage. The animal lives in this wheel. The numbers of revolutions which it will make in a given time-interval are found to vary with: time of day, time of feeding, sex, age, and the like. Young rats are far more active than old. Mature females have a pronounced four-day activity cycle (corresponding to their four-day œstrus cycle), and so on.

The *general exploratory* drive has been measured by numbers of crossings of the electric grill which the rat will make in order to reach a small complicated maze, especially designed by Dashiell to encourage exploration. It is found that a rat will cross the grill to reach such a to-be-explored maze slightly more frequently than he will cross it to reach a simple empty box.

The *rest* and *sleep drive* and the *urination* and *defecation drive* have not as yet been subjected to controlled investigation.

The *play* and the *æsthetic drives.* The beginnings of these drives, as we know them in man, are probably not to be found in the rat, but they certainly are to be found in the chimpanzee. Köhler in describing the animals observed by him at the experimental station on the Island of Tenerife in 1913–1917 cites various examples. I quote a couple of his descriptions. The first has to do with a primitive sort of dance. He writes:

> The whole *group* of chimpanzees sometimes combined in . . . elaborate and semi-rhythmic *motion-patterns.* For instance, two would wrestle and tumble about near some post; their movements would become regular and tend to describe a circle round the post. . . . One after another, the rest of the group approach, join the two, and finally they march in an orderly fashion and in single file round and round the post. Their movements become animated; they no longer walk, they trot, and as a rule with special emphasis on one foot, while the other steps lightly; thus a rough approximate rhythm develops and they tend "to keep time" with one another. They wag their heads in time to the steps of their "dance" and appear full of eager enjoyment of their primitive game.[1]

Or, again, he describes a case of primitive painting:

> Once, when lumps of white clay were brought to the playground, the animals suddenly began to paint, without any stimulation, and whenever afterwards they again got clay, the same game would begin. At first . . . (they) licked the unknown stuff; . . . . The result being unsatisfactory . . . they wiped their protruding lips on the nearest object they could find, and, of course, made it white. But, after a while, the painting of beams, iron bars, and walls grew to be quite a game of its own; the animals would seize the clay with their lips, sometimes crush it into paste in their mouths, moistening it, and would then apply the mixture, make fresh paint, and daub again, and so on. *The point is the painting, not the chewing of the clay;* for the painter himself, and the rest of them, when not too much occupied with their own affairs, are most interested in the result.[2]

---

[1] W. Köhler, *The Mentality of Apes.* N. Y.: Harcourt, Brace and Company, 1927, p. 314f.

[2] *Op. cit.,* p. 96f.

These two quotations present examples of something which might be called either play or the æsthetic impulse. In fact, these two drives would seem to be very closely allied and for the purposes of the present discussion we shall not try now to distinguish further between them.

To sum up for all the appetites, it would appear that each is set in motion by some peculiar internal metabolic condition (state of the breasts, hunger, thirst, sex, need for exercise, and the like). This metabolic condition occurs in apparently more or less regular cycles due to combinations of internal and external conditions. And when it is in force the animal is *driven* until an appropriate consummatory object is found in the presence of which latter a corresponding consummatory response occurs. This consummatory response then relieves the internal metabolic condition and produces, at least temporarily, a final complementary state of internal *satiation*.

(ii) Consider, now, the *aversions*. They have not been subjected to any precise experimentation. I shall suggest the following list for human beings:

The Aversions

Cold-avoidance
Heat-avoidance
Danger-avoidance (*i.e.*, Fright)
Obstruction-avoidance
(*i.e.*, Aggression)

Each of these aversions is set off, not by an internal metabolic condition (as is an appetite), but by an *evoking environmental object* (or situation), *i.e.*, cold, heat, danger, or obstruction. Further, whereas an appetite was seen to be primarily a *getting-to* a final internal state of "*satiation*," each of these aversions is, rather, a *getting-from* an internal state of "*sufferance*." For cold, heat, danger, obstruction tend to produce internal sufferances. And it is such internal sufferances which are the terminal situations ultimately being got from in each case.

This is about all we can now say about the aversions. Obviously a lot more work has to be done on them.

(*B*) *The Social Drives*

Turn to the social drives. It is primarily because of Köhler's early observations made at Tenerife and because of the long series of more recent and more experimentally controlled studies carried out at Yale under Yerkes' direction that one is driven to believe that there really is a set of innate social drives in the chimpanzees and hence, presumably, in man also. As a very tentative list of such social drives I would suggest the following:

Social Drives

Gregariousness	*i.e.,*	Returning to company of others of the group
Loyalty to group	*i.e.,*	Defending other members of the group against attack
Imitativeness	*i.e.,*	Copying actions performed by other members of group
Dominance	*i.e.,*	Dominating over another individual
Submission	*i.e.,*	Submitting to another
Competitive acquisition	*i.e.,*	Piling up material for the future—such activity being enhanced by the presence of other individuals
Sharing with and soliciting from others	*i.e.,*	Giving to another individual. Receiving from another individual
Coöperation-tendencies	*i.e.,*	Working with another individual for a common goal

This list is not to be taken too seriously. As it stands, it corresponds more to a mere concrete series of observations and experiments, which happen to have been carried out thus far, than to any final, and careful, analysis of exact interrelationships. Further, this does not permit the citation of much of the actual work which has been done. I shall limit myself to two accounts: a description of "Loyalty to the Group" presented by Köhler and an investigation of "Sharing with others" carried out by Nissen and Crawford at Yale.

First for *Loyalty to the Group,* Köhler writes:

> If one chimpanzee is attacked before the eyes of the group, great excitement goes through the whole group. It will happen that, under the influence of the climate, one punishes a wrongdoer with a heavy

> blow. The moment one's hand falls on him, the whole group sets up a howl, as if with one voice. . . .
>
> When the apes have grown much older . . . and especially after they have arrived at sex maturity, I find the drive of the group to repulse an assault on one of its members grown inordinately stronger. In the end, one has to give up punishing even bad offenses, when the whole group is in the same room as the wrongdoer. At times the most insignificant episode between man and an ape, which arouses the ape to cry in anger against the enemy and spring against him, is sufficient for a wave of fury to go through the troop; from all sides they hurry to a joint attack. In the sudden transfer of the cry of fury to all the animals, whereby they seem to incite one another to ever more violent raving, there is a demonic strength, coming, surely, from the very roots of the organism. It is strange how convincing, one might say full of moral indignation, this howling of the attacking sounds to the ear of man:[3]

Now for the investigation of "Sharing with others" made by Nissen and Crawford. Pairs of animals were put together in adjacent cages. One animal in each pair was given food or food-tokens in the nature of poker chips (which the animals had learned previously to use in penny-in-the-slot food-vending machines). And observations were made of the readiness of the one animal which received the food (or the mere token) to share with the other animal. The results were of the sort indicated by such titles of photographs as the following:

"Bula begs food unsuccessfully"
"Bimba gives Bula token"
"Bula shows interest in onion but does not
attempt to take away"
"Bimba teasing Bula"
"Bula passes a token"
"Bimba gives Bula tokens"

Finally, to quote from a summary by the authors:

> It appears that food-sharing as we have seen it among chimpanzees, when compared with other forms of mutual aid or social service found in animals, is relatively variable, non-compulsive in character, adaptable to varying conditions, not strictly circumscribed by physiological or external factors but influenced strongly by previously established social relationships.[4]

---

[3] W. Köhler, *The Mentality of Apes.* N. Y.: Harcourt, Brace and Company, 1927, p. 286f.

[4] H. W. Nissen and M. P. Crawford, A preliminary study of food-sharing behavior in young chimpanzees. *J. Comp. Psychol.*, 1936, 22, p. 415.

This completes my very inadequate account of all three kinds of drives—the simple appetites, the simple aversions, and the social drives.

Appetites are to be defined as gettings-to internal satiations, aversions as gettings-from internal sufferances. And social drives I would define as both gettings-from sufferances and gettings-to complementary satiations—for example, getting from the sufferance produced by the begging of a partner and getting to the satiation produced by the resulting contentment of that partner.

We are now ready to turn to our second main concept, that of values and valences.

2. Values and Valences

When any one of the three types of drive is in force, we may assume a state of internal tension. In the case of an appetite, this internal tension drives the organism until an appropriate environmental consummatory object is reached, whereupon a consummatory response follows, a state of satiation results, and the tension subsides. In the case of an aversion, the tension drives the organism until an environmental evoking object and a thereby threatened state of sufferance are both got from, whereupon the tension subsides. In the case of a social drive, the tension drives the individual until an environmental evoking situation (and its accompanying sufferance) are got from and a complementary environmental consummatory situation (and its accompanying satiation) are got to, whereupon, again, the tension subsides.

The point, I would now make, is that any one of these driving tensions, while it is in force, is to be said to endow the environmental consummatory or evoking situations with value—consummatory situations with positive value and evoking situations with negative value. But such positive and negative values, thus inhering in environmental consummatory and evoking situations, are *instrumental* only. Final, *intrinsic*, values inhere nowhere but in the internal states of satiation and of sufferance. In the case of hunger, final intrinsic positive value inheres only in the final state of hunger-satiation. In the case of cold-avoidance, intrinsic negative value inheres only in the state of internal sufferance which would result from the external cold.

And in the case of "group loyalty" intrinsic negative value inheres only in the internal sufferance which accompanies not fighting for one's group and intrinsic positive value inheres only in the internal satiation which accompanies fighting for one's group.

Instrumental values, however, are attached not only to the final environmental consummatory and evoking situations but also to ordinarily neutral environmental objects (or situations) in so far as these latter came to serve as consistent *means* or *hindrances* towards the consummatory, situations or away from the evoking situations. For example, a restaurant has instrumental *positive* value because it is a means for getting to food. And poverty has instrumental *negative* value because it is a hindrance both in getting to food and in getting away from cold and from a lot of other sufferances besides.

To sum up, then, it is the internal states of satiation and sufferance which alone are endowed with *intrinsic values*. Environmental objects and situations, whether terminal (*i.e.*, consummatory or evoking) or merely means and hindrances, are endowed with mere instrumental values.

But it now appears that we also need a second concept in addition to that of value and I shall use the term "valence" for this further concept. The term, "valence," is the translation which has been suggested for Lewin's original German term "*Aufforderungscharakter.*" The literal meaning of the latter is "invitation-character." And this concept of "invitation-character" or "valence," we shall now find helpful to refer to values, not as they actually are, but as they are for the given individual in question. The "values" are what they really are, as seen by us the *omniscient observers*. The "valences," on the other hand, are those values which the given individual himself responds to. An individual's valences will tend only in part to correspond to what *we* know to be the true values. For example, a given food may be known by us, the omniscient observers, to have decided positive value for a given hungry individual. But to him (as we can tell from his behavior) it may have practically *no* positive valence. He "ought" to be attracted by it, but he isn't. Or, again, a given building distant on the highway because it contains a restaurant, will be said by us to have considerable positive value for a hungry autoist.

But to this autoist himself, it may have practically no positive valence. He may never have seen the building before and hence not "believe" that it contains a restaurant.

If we attempt now, very briefly, a general survey of all such possible types of divergence between valences and values, it will appear, first, that the cases in which an individual fails to experience positive or negative valences corresponding to *final consummatory* and evoking values are probably few. A man is not very apt to overlook either actual food or actual cold. Secondly, however, there will be a relatively large number of cases in which an individual experiences means or hindrance valences (positive or negative) when there are not any corresponding values. A miser's love of money is the stock example. The miser experiences a positive *valence* which far exceeds any corresponding real positive *value*. And the phobias, *e.g.*, claustrophobia and agoraphobia, are other instances. They are cases of strong negative *valences* for types of situation the real negative *values* of which are practically nil.

It will appear, however, that to consider intelligently all such types of divergence between the actually experienced valences and the true values we shall need to consider first another concept—that, namely, of "beliefs." For it is the individual's beliefs (true or erroneous) which finally determine in how far environmental objects take on for him the valences which they do.

Let us turn then, now, to beliefs.

3. Beliefs

A distinction is to be made first between *instrumental beliefs* and *equivalence beliefs*. Instrumental beliefs are to be defined as the readinesses of an individual to accept certain objects or situations as either means or hindrances, or signs (see below) relative to other objects or situations. And equivalence beliefs are to be defined as the readinesses of an individual to accept certain objects or situations as *functionally equivalent* to other objects and situations. Let me illustrate in terms of the two types of classical learning experiment.

The first of these is the type of experiment originated by Thorndike in his early investigations of the hungry cat getting out of a cage. It is know as trial and error learning. Shut up with food outside, the animal learned (through repeated so-called *trial* and *error*) that attacking the bars did not get her out—that the bars were, in short, hindrance-objects. She also learned by this same trial and error that pulling at a particular loop of string, hanging down in the middle of the cage, did get her out, *i.e.,* that such a loop was a positive means-object. As a result, she became ready thereafter (*i.e.,* after a certain number of trials) *not* to attack bars and *to* pull at the loop. Or, in our present terminology (which by the way, would no doubt be very shocking to Thorndike) she acquired *instrumental beliefs* to the effect that bars are negative (or hindrance) objects and that loops are positive (or means) objects. And hence, further, the bars took on for her negative instrumental valences and the loop took on a positive instrumental valence. Furthermore, in this particular case, these her acquired valences were correct. They corresponded, that is, to what we, the omniscient observers, can see to have been the actual true values.

But such a type of learning also establishes *equivalence beliefs.* For, once the cat had learned *not* to attack the bars and *to* attack the loop, Thorndike introduced transfer experiments in which the animal was then tried out in other somewhat similar cages. And it was found that now other similar bars would also not be responded to (that is, would be accepted by the cat as equivalent to the original bars) and other opening devices relatively similar to the original loop *would* be responded to (that is, would be accepted by the cat as equivalent to the original loop). But, if such new bars and new opening mechanisms were too divergent from the original ones, the learning, that is, these, her acquired equivalence beliefs, would not carry over.

Any case of trial and error learning may be said, then, to result in both instrumental and equivalence beliefs. The former determine what positive and negative instrumental valences have come to adhere to what given specific features in the original set-up. And the latter, the equivalence beliefs, determine to how wide a range of other, more or less similar, objects these same instrumental valences will carry over.

Consider now also the second classical type of learning experiment—the conditioned reflex experiment. You are all familiar with this as originally carried out by Pavlov. A dog had a salivary fistula made, whereby the saliva he secreted on one side came out through a tube cemented to the cheek and could be measured. A metronome was sounded prior to, and continuing simultaneously with, the presentation of food. After a few trials, the hungry dog began to secrete saliva at the sound of the metronome, just before the food itself actually appeared. Now such learning also involves the acquisition of an instrumental belief—that is, a special variety of instrumental belief which I would call a *sign belief*. The animal learns, namely, to take the metronome as a *sign* for the coming of food. And secondly this type of learning also involves an equivalence belief. For it will be found that this conditioned secretion will occur for quite a range of metronome rates other than the original one. Such other metronome rates are, in short, believed by the dog to be equivalent to the original one. And, it may be noted in passing that further experiments can be carried out in which such equivalence beliefs as to signs can also be narrowed down by special training.

Finally, I would say that ordinary learning in man also results in just such instrumental and equivalence beliefs as those which get set up in cats and dogs.

The next important point to note is that in both man and the lower animals, when the situation changes, the normal thing is for new learning to occur. If the means, hindrance and sign relationships in the environment be changed, then the man or the lower animal will normally relearn and acquire the necessary new beliefs. And in all these cases where such new learning is readily possible I would call the original beliefs—*rational*. But, as we shall also see in a moment, there are many cases in which the first beliefs cannot thus be readily given up in favor of the necessary new ones. In these latter cases I shall call the original, too strongly held to, beliefs *irrational* or *traumatic*.

Our next question is, then, what are the types of situation which tend to give rise to irrational or traumatic beliefs? And my answer will be that they are situations which involve too much frustration—*i.e.*, situations in which a real solution, a real learning, is relatively difficult and beyond the capacity of the organism. Such situations occur, of course, very frequently in childhood. And I would suggest in passing that it was one of Freud's major contributions to have discovered the importance of such early childhood frustrations and the fact of the consequent irrational, or traumatically held to, beliefs which as a result tend to get established in childhood and to persist in hidden forms throughout later adult life. But let us consider now a few of the characteristic types of such irrational, traumatically held to beliefs. We may call them frustration mechanisms.

4. The Frustration Mechanisms

The following list is a very abbreviated one. Actually Freud, and others, have uncovered some twenty or more such mechanisms. This list will suffice, however, for illustrative purposes.

Frustration Mechanisms

- Fixation
- Symbol-Magic
- Displaced Aggression
- Phantasy and Dreams
- Sex Perversions or Sublimations
- Going out of the field, Suppression and Regression
- Identification
- Vicarious Identification
- Group Identification

These are all names for characteristic ways in which an organism, faced with a relatively difficult problem of learning the appropriate signs and means for getting on to some final satiation or away from some final sufferance may tend to get caught by erroneous, and irrationally held to, beliefs concerning such signs and means. Further, it will appear that these erroneous and irrationally held to beliefs are, in the last analysis, *equivalence beliefs.* And they may be divided into the three subclasses of: (1) an overready acceptance of a *frequent means or sign* as *functionally equivalent* to that for which it is the means or sign; (2) an overready acceptance of a *similar* as *functionally equivalent* to that to which it is similar; and (3) an overready acceptance of a *contiguous accompaniment* as *functionally equivalent* to that which it accompanies. But let us consider this list of mechanisms, now, one at a time.

Fixation. This term is used to designate a means which was originally experienced many times over but which continues to be irrationally held to after it has ceased to be correct. For example, a man who at the age of four was properly in love with his mother continues, as an adult, to be "fixated" on her. As a result he fails to marry; and he finds normal sexual relations with women of his own age and class difficult. This, of course, is the famous Œdipus Complex.

But fixations are also to be found in less poetical areas than "Hamlet" and the "Œdipus Rex." Even the white rat may exhibit them. If a rat be taught first to choose the right-hand path and then the food be switched to the left-hand path, it may take the beast an undue number of repetitions to learn the new direction, if the first one were very much overlearned. The over-frequent exercise on the right will have caused him to become fixated upon that side—*especially so,* if in the original set-up his frustrations were great—if he received an electric shock at the choice point, or if he had especial difficulty in learning it; or if he were especially hungry.[5]

That is, all such fixations, whether of the man on his mother or of the rat on its path, are to be characterized as erroneous and irrationally held to equivalence beliefs. The mother, a mere early and frequent means-object, is held on to throughout life, irrationally and erroneously, as *equivalent* to the actual satiations which her original caresses produced. And, similarly, the right-hand path is erroneously and traumatically held on to by the rat as *equivalent* to the hunger-satiation which it originally led to.

Turn to the next item.

Symbol-Magic is a concept originally emphasized by Kempf in this country. It describes the case in which, when a goal is debarred, we resort to some sort of hocus-pocus which simulates that goal. Thus, *hanging* or *burning in effigy* is the classical example. The real goal, the persecution and death of some hated character, cannot be got to so we indulge in the symbolic acts of burning or hanging a likeness of that hated character. A similar, or symbol, is accepted as functionally equivalent to that to which it is a similar.

Displaced Aggression.—Normal aggression is an appropriate attack upon a true barrier. Displaced aggression, in contrast, is our name for the case where, because an aggressive attack upon the true barrier itself is unsuccessful, or prevented, the individual takes it out by an attack upon some innocent bystander or upon some inanimate object. From a recent issue of *Life* it appears that a couple of young modelers have capitalized on this propensity by inventing a small clay figure known as a "wackeroo." When difficulties arise one grabs the wackeroo and hurls it across the room with a satisfying crash—obviously, a grand case of displaced aggression. But such displaced aggression is also but another case of a traumatically held to and erroneous equivalence belief. It is, in short, the belief that the destruction of a mere accompaniment of a barrier is functionally equivalent to the destruction of the actual barrier itself.[6]

---

[5] See experiments by Ellis and Hamilton, by Gilhousen, by Elliott, by Krechevsky and Honzik, by Hamilton and Krechevsky, and by Sanders.

Phantasy and *Dreams* deserve a longer discussion than I have time for. They can, however, be characterized, very briefly, as cases in which a combination of similars and appropriate means are accepted as functionally equivalent to a distant, relatively unattainable goal. In phantasy one mentally rehearses the goal. And such a mental rehearsal is actually part of the total appropriate means necessary for getting to the goal. It is a pure phantasy, however, when instead of being a mere planning for, this mental rehearsal comes to be accepted as functionally equivalent to, and a substitute for, the true goal itself.

Dreams are rather similar. The distinctive feature about them is, however (as Freud demonstrated), that in the dream the real meaning is apt to be disguised. Dreams occur when there is such a strong barrier barring off the real goal (in this case an inner barrier set up by early moral training) that the goal not only cannot be reached but also cannot even be thought about in undisguised form. A dream is thus a very absurd, as well as an irrationally held to and erroneous, equivalence belief. It is a belief to the effect that thinking about a goal in symbolic terms is functionally equivalent to actually attaining that goal.

Sex Perversions and Sublimations.—These have been grouped together because in both of them the true sex goal is debarred and obvious substitutes are accepted. These substitutes may be preliminary types of means-activity or merely symbolic activities. The big difference between perversions and sublimations is that in sublimations the accepted substitute is very far removed from the original sex activity in appearance and it is something which the group declares to be socially useful.

---

[6] A recent cooperative study at Yale by Dollard, Doob, Miller, Mowrer and Sears has contributed greatly to our knowledge concerning such Displaced Aggressions.

Going Out of the Field, Suppression and Regression.—The concept of "Going out of the Field" was first hit upon by Dembo in Lewin's laboratory. Miss Dembo's subjects were presented with unsolvable tasks which, however, seemed solvable. And she found that as time passed and a subject continued not to find a solution he or she would experience tremendous frustration. As a result some of her subjects refused to continue the prescribed task, turned to other activities, or even ran bodily out of the room. For all such behaviors she coined the comprehensive term "Going Out of the Field." Such "Goings Out of the Field" were all cases in which the negativity of the barrier may be said to have spread erroneously and irrationally to the goal behind that barrier. It is the old sour grapes phenomenon. In extreme frustration the subject falls into the incorrect and traumatically held to "equivalence belief" that because the barrier is bad the goal behind, and contiguous to, that barrier is also bad.

Regression and Suppression I class with "Going Out of the Field" because they also are cases where a positive goal is abandoned because of the negativity of the barrier in front of it. By regression we mean the case where the individual gives up the immediate adult goal altogether and goes back (regresses) to some childhood goal. And by suppression we would designate the case where such a giving up, *i.e.,* going out of the field of a goal, does not really get rid of the goal altogether but merely pushes it away from immediate direct attention and allows it to continue to exert an influence in indirect ways.

Identification is easily observed in young children. A child (faced with the complexities of its environment) tends to "identify" with the parent, or with some other adult, of the same sex. It tries to copy what it takes to be this adult's aims or goals. According to Freud, this phenomenon occurs especially during the early Œdipus stage. At the age of four or five the little boy wants to replace the father in his mother's affections and so tries to become like the father. Or the little girl wishes to replace the mother in the father's affections and so tries to become like the mother. But in any case, whether or not we accept this supposed causal importance of the Œdipus stage, we cannot deny that there are such tendencies for children to identify with adults and usually, though not always, with adults of the same sex as their own. And, finally, it appears that such identifications also are to be classed as erroneous and irrationally held to *equivalence beliefs.* The child, irrationally and erroneously, believes that the adult's goals are functionally equivalent to, and are to be taken as substitutes for, his own.

The above may be called simple identification. I wish now, however, to suggest a further related phenomenon to be called *Vicarious Identification.* I would mean by it the fact that the child not only tries to copy the adult's goals but will also, in some degree, come to accept the latter's successes *vicariously* as substitutes for his own successes. The father's prestige will be accepted by the boy as a substitute for his own prestige. And the mother's social triumph will be accepted by the daughter as her social triumph. Obviously, here again is a case of an erroneous and traumatically held to equivalence belief.

Finally, there is still a third related phenomenon which I would call *Group Identification.* A child, or indeed an adult will, when frustrated, identify not only with parents or leaders but also with the whole group such as the family, the school, the economic class, or the nation. Such group identifications with family or larger group are associated with identifications with the parents or leaders of the group but also contain an element of identification with the whole group itself as some sort of a mystical unity. The individual comes to accept the success of this group vicariously in default of any success of his own. Again it is a case of an erroneous, and traumatically held to, belief in the functional equivalence of such successes of one's family, class, nation or, indeed, of one's "hemisphere" (and also of the leaders thereof) as identical with and hence substitutable for one's own successes. The world's history is certainly all too full of such instances in which the common man has sacrificed himself quite erroneously for the glory and prestige of his group. And yet such group identifications are today still prevalent.

Why do we not learn otherwise? Two answers may be suggested. First, such apparent sacrifice is sometimes actually valid. Sometimes the successes of one's group really are conducive to one's own ultimate biological and social satiations and non-sufferances. And, second, this group-identification tendency is also at the same time reinforced by, or a direct expression of, one of the basic social drives, namely, that of *"loyalty to the group."* You remember Köhler's chimpanzees who were all ready to attack the experimenter and be punished themselves when any one member of their group was punished. In short, the sacrifice of oneself for the group (even though this may interfere with some or even all of one's own biological and social satiations) provides at the same time direct satiation for *one* of the social drives. In a word, although identification with the group is in part a mere irrationally held to and often erroneous *equivalence belief,* it is also in part the direct expression of one of the simple instinctive drives. Hence its everlasting ubiquity.

But let me sum up, now, as regards all these frustration mechanisms. They occur primarily in situations of too great frustration. And in essence, they are beliefs (usually both erroneous, and irrationally held to) that "means," "similars," or "contiguous accompaniments" are the same as (*i.e.*, functionally equivalent to) the other objects or situations to which they are mere means, similars, or accompaniments. And the valences, positive or negative, which belong properly to such other objects get transferred to these their means, similars or accompaniments. Further, such beliefs, being irrational or traumatic as well as erroneous, require excessive amounts of contradicting experience to be overcome. In fact, it is usually necessary to get the help of some second person. The function of a clinician—psychoanalysist or mere eclectic—may, indeed, be said to be primarily that of helping the patient to substitute, for such erroneous, and traumatically held to, *equivalence beliefs,* more valid, and more rationally held to, *instrumental beliefs.* For it is in this way that the valences of the patient will come more and more to correspond to the true values which (given his drives and capacities) his environment actually holds out for him.

This completes my very inadequate indication of some of the problems of motivation, learning and adjustment in the individual as such.

Group Motivation and Adjustment

Before concluding, however, I feel that I must hazard a word as to certain problems of group motivation and adjustment. For today the problems of the individual recede in importance. What boots it whether you and I achieve satiations or suffer sufferances, if our whole society, our whole civilization is to go under? I have but little helpful or perhaps true to say but I would nonetheless like to suggest a couple of inferences from the arguments presented above. For it appears to me that two at least of the above mechanisms: that of *Group Identification* and that of *Displaced Aggression*—are much to the point. The Nazis and the Facisti today obviously present, it seems to me, astounding examples of intense *Group Identifications* and of virulent *Displaced Aggressions.*

In Germany, Group Identification—the traumatic belief that the successes of the state are substitutable for the successes of the individual—has reached extraordinary proportions. Long before the last war this belief was officially accepted and taught. Further, as a result of the frustrations of the individual coincident upon the Treaty of Versailles and also coincident upon a long accepted relatively Spartan system of childhood training, Germany has also long been a nation of frustrated individuals—ones full of aggressions ready to be displaced. The strong identification with the state and these dammed up aggressions inevitably combined in an attack of overwhelming fury against supposed enemies of the state. The only cure for Germany would be a less rigid and Spartan system of training and a less mystical belief in German "Blut und Boden."

Italy presents a similar though less extreme picture. The youthful frustrations, due to types of family and school discipline, are probably not as great in Italy as in Germany, nor were the Italians in general so frustrated by the outcome of the last war. In Italy the "frustrations" come more from the actual poverty of the country. And, perhaps the Italian language does not foster traumatic identifications with the group as easily as does the German. In any case, Fascist Italy, though like Nazi Germany, is less so and less successfully so.

But turn now, finally, to ourselves. The significance of our Democracy has been, in the past, both the ever repeated breaking down of tendencies toward traumatic identifications with the state or its rulers, and the ever repeated prevention of too great frustrations of the individual. By our free family life, by our non-authoritative educational system, and by our democratic political procedures (together with the grace of an open frontier plus two oceans) we have insured both that the individual did not come to believe in the welfare of the group (or of its leaders) as more important than his own, and also that our non-satiations and our sufferances as individuals should not become too great and not build up into too many displaced aggressions. Such *has* been our democracy.

But today all is threatened. We Americans, like the denizens of the Old World, are beginning to talk primarily about group survival, group success, group enemies. The questions therefore which today we must ask ourselves are: How far are we (like the Germans and the Italians) merely traumatically identifying with our group? And how far have the recent depressions and recessions, the horrors of unemployment and all our consequent frustrations merely led us into displacing our aggressions against enemies abroad rather than directing them against the true barriers at home? Or, how far, on the contrary, *are* we validly responding in both matters to real instrumental relationships? How far, that is, *would* a German victory really mean the end of America? How far would it truly mean the greatest of sufferances for all of us as individuals?

One does not know. History alone could provide the true answer. Some of our present beliefs and consequent valences certainly are traumatic. Some of them are no more than the purely irrational belief in the mutual "functional equivalence" of all people who speak English or of all people who live in the Western Hemisphere. And some of our present activity is certainly no more than a mere matter of us old men displacing our aggressions and being ready to send our boys to do our killing for us. But this is not the only side of the picture. For, even if, in truth, our present driving valences do come largely from traumatic group identifications and traumatic displaced aggressions, these valences are now so all-powerful in us that we cannot protect ourselves against them. And, if on the other hand, these driving valences are in large measure the expression of true, and rationally held to, instrumental beliefs as to what really would happen, we do not wish to be protected from them. There are, then, as I see it, but three commitments now left to us.

One. We must so work that some larger, more powerful, more international group shall emerge from this war, a group with which our children and our children's children can more happily and more safely identify. For, if no such group, no honest "League," "Federation," or "Union Now," be born this time either, then God help our grandchildren twenty-five years hence when we and they shall have invented still more powerful engines of destruction.

Two. We must so work throughout this very period of emergency as to insure that our identifications with our group—the sacrifices that we ask of ourselves, of the young and of labor—shall be as realistic and as untraumatically held to as possible. We must see to it that these identifications and these sacrifices have a different, less traumatic, coloring from that which they do in the very groups whom we are opposing.

Lastly. We must so work as ever to remember that in the end it is only the satiations and non-sufferances of individuals—the final satisfactions of the biological and social drives which alone can make a good and worthy society. A surviving America with the same old soil erosion, pellagra, discrimination against racial minorities, and unemployment will not have been greatly worth the throwing away of our steel, our oil and our sons.

DETERMINATION OF LIMB DARKENING IN ECLIPSING BINARIES FROM COLOR-INDEX OBSERVATIONS

ZDENĚK KOPAL

Harvard College Observatory

Abstract

After a brief survey of the problem (Sections 1 and 2), a method is developed for determining limb-darkening coefficients from the light curves of eclipsing binaries in minima observed in two or more wave lengths applicable to both partial and total eclipses. For the purpose of a least-squares solution, the equations of condition are reduced to standard linear form (Section 3). Several alternative forms of the equations of condition are suggested, and their weights and the effects of errors in the photometric measurements or geometrical elements are discussed (Sections 4 and 5). It is pointed out that, in the case of total eclipses, moderately precise measurements of the color indices can yield a relatively accurate value for the difference in the darkening coefficients for the two respective wave lengths.

Application is then made to U Sagittae and U Cephei as illustrative examples. In Sections 7 and 8 the geometrical elements and period of U Sagittae are re-investigated; Section 9 contains a discussion of the observations of U Sagittae (spectrograms taken in different partial phases) and their reduction; and in Section 10 the individual darkening coefficients as well as their difference are derived for λλ4050 and 4750. The resulting values are: λ4050, $u = 0.5 \pm 0.2$; λ4750, $u' = 0.4 \pm 0.2$ (p.e.); $u - u' = 0.09 \pm 0.06$ (p.e.). Within this interval in wave length, the limb darkening of the B9 component of U Sagittae thus seems to be nearly constant, or increases but slowly with increasing frequency of light. In Section 10, Rosenberg's color-index measurements of U Cephei are re-discussed, and found to be insufficiently accurate for quantitative conclusions regarding the limb darkening of its bright component. Finally (Section 11), previous results for other stars are compared with those obtained here.

I. INTRODUCTION

1.—As is well known, measurements of limb darkening provide us with valuable information concerning the physical processes prevailing in stellar atmospheres. This phenomenon is due to the fact that the radiation which emerges from an absorbing atmosphere originates, on the average, at a greater optical depth, and therefore corresponds to a higher temperature, if viewed normally to the surface than if viewed tangentially. The apparent surface temperature will thus depend on the angle of foreshortening, being highest at the center and gradually decreasing toward the edge. As a consequence, the surface brightness behaves likewise. The intensity of radiation emitted at a point where the surface-normal makes an angle θ with the line of sight is given by

$$I = \int_0^\infty E(\tau)e^{-\tau \sec \theta} \sec \theta \, d\tau, \tag{1}$$

where τ denotes the optical depth. If we make the usual approximation that $E(\tau)$ is proportional to τ, the integration gives

$$I = I_0(1 - u + u \cos \theta), \tag{2}$$

where u, the coefficient of limb darkening, is

$$u = \frac{1}{1 + \frac{8}{3}\frac{\lambda T}{c_2}\left(1 - e^{-c_2/\lambda T}\right)\frac{\kappa_\lambda}{\bar{\kappa}}} . \tag{2.1}$$

Here c_2 is Planck's second constant (1.432 cm deg); T, the surface temperature connected with the effective temperature T_e by the relation

$$(T_e/T)^4 = 4/\sqrt{3};$$

and κ_λ is the continuous absorption coefficient for the wave length λ, while $\bar{\kappa}$ is the opacity coefficient (Rosseland mean). Hence u proves to be dependent on the temperature and on the frequency of light, as well as on the nature of the absorption processes in stellar atmospheres. Its values would seem to be predictable on a theoretical basis, were it not for the fact that the hypothesis that $E(\tau)$ is linearly proportional to τ, which we employed in the integration of (1), is obviously but an approximation and, in addition, the constant of proportionality depends on the quality of the stellar radiation. Moreover, the values of $\kappa_\lambda/\bar{\kappa}$ are a difficult subject for rigorous calculation. Pannekoek[1] and others have endeavored to calculate them as a function of the temperature for different wave lengths and surface gravities; yet, despite a great amount of work, our positive knowledge is still far from satisfactory. Hence the problem of finding the degree of darkening for a star of given temperature and effective wave length is semi-empirical: whenever the value of u can be inferred from the observations, we gain a valuable indication of the correctness of the theoretical predictions.

2.—Apart from the sun, eclipsing binaries are the only possible source of information on this parameter. It has been pointed out by Russell and Shapley[2] that limb darkening profoundly influences the determination of the geometrical elements of eclipsing variables from their light curves. But the simultaneous determination of limb darkening with the geometrical elements has proved to be an exceedingly difficult task, because the effects of limb darkening are largely correlated with those produced by various other geometrical elements. In order to separate them, observational data of the utmost accuracy are required. Numerous earlier attempts at the quantitative determination of limb darkening from moderately accurate light curves necessarily failed, or, at best, were able merely to indicate which of the two limiting solutions, "uniform" or "completely darkened," is nearer to reality.[3] Pannekoek and Miss van Dien [4] applied the method of least squares for adjusting the arbitrary constants and were the first to realize fully the intrinsic difficulty of the problem. Not until quite recently have such investigations led to positive results. Wyse [5] has improved the method of simultaneous determination of the darkening coefficient and other geometrical elements by least-squares corrections; and Kron[6] has finally produced such an accurate light curve of *YZ* Cassiopeiae that he has succeeded in deriving the limb darkening coefficient of its brighter component with a probable error of 8 per cent.

---

[1] *Amsterdam Publ.* 4, 1935; *cf.* also *M.N.*, **95**, 733, 1935.

[2] *Ap. J.*, **36**, 239, 1912.

A different course may be pursued if observations in different wave lengths are available. Limb darkening manifests itself not only by decreasing the total intensity toward the edge, but also by making the limb redder than the center. Consequently, as the eclipse progresses, the color-index of the visible crescent must be subject to gradual changes. The color index of an eclipsing system at any one moment is naturally the mean of the color indices of the two components, weighted according to their luminosities. If the eclipsed disc is "uniform," its luminosity during the eclipse decreases proportionally to the hidden area; and if the temperatures and relative dimensions of both components are known, the color index of the system at any specific moment can be predicted as a function of the phase angle. Systematic deviations from the predicted relationship are definite indications of limb darkening. While the first method, based on observations in one wave length only, can preferably be applied to systems with annular deep minima, the latter method will prove to be advantageous for systems the deep minima of which are total. Both methods require highly precise observations if the individual coefficients of darkening are to be obtained; however, even moderately accurate measurements of the color indices can yield a relatively accurate difference between the darkening coefficients in the two respective wave lengths.

---

[3] For instance, Stebbins (*Washb. Publ.* 15, 41, 1928) on α Coronae Borealis; or McDiarmid (*Princ. Contr.* 7, 50, 1924) on TX Cassiopeiae.

[4] *B. A. N.*, **8**, 141, 1937.

[5] Lick Bull. 494, 1939.

[6] Lick Bull. 499, 1939.

The first and so far the only investigation of the second kind is that of U Cephei by Rosenberg,[7] who carried out the photographic photometry of this system in two widely separated effective wave lengths. Rosenberg, however, failed to investigate adequately the nature of the problem, its determinateness, or the effects of errors entering through the underlying observational data. Moreover, it can be shown that if the geometrical elements are known from previous investigations, a knowledge of the complete light curves (*i.e.*, of instantaneous brightness) of a system in two effective wave lengths is superfluous; the only necessary requirement is the knowledge of the color-index variation between minimum and full light. In Part II of the present paper an exhaustive analysis of the problem will be given. Part III will contain an application to U Sagittae, based on the measurements of spectrograms taken in different partial phases; and in Part IV Rosenberg's data on U Cephei will be subjected to a similar re-discussion. Finally, in Part V, will appear a general discussion and comparison of the results arrived at below with those obtained previously for other stars.

II. THEORY OF THE EFFECT

3.—Let l denote the relative luminosity of an eclipsing system at phase angle ψ (supposed within eclipse), and let α be the fractional loss of light of the eclipsed disc of a component which, for convenience, we shall call the primary. Then, at any phase,

$$l = L_2 + L_1(1 - \alpha), \tag{3}$$

where L_1 and L_2 are the relative luminosities of the primary and secondary component, respectively. If observations in two wave lengths are available, and if l, l' denote the luminosities of the system in any two particular wave lengths λ, λ', we obtain by division

---

[7] *Ap. J.*, **83**, 67, 1936.

$$\frac{l'}{l} = \frac{L_2'}{L_2} \cdot \frac{1 + (1 - \alpha')\Lambda'}{1 + (1 - \alpha)\Lambda}, \tag{4}$$

where

$$\Lambda = \frac{L_1}{L_2};$$

i.e., an equation governing the variation of the intensity gradient with the phase. The value of the left-hand term could readily be derived from the observations, but accurate measurements of such a nature are known to be difficult. Much easier to measure is a quantity $S_{\lambda\lambda'}(l'/l)$, where $S_{\lambda\lambda'}$ denotes effects which depend on the wave length (such as atmospheric and instrumental absorption, plate sensitivity, etc.), but are constant as long as we use the same instrumental arrangement, check the irregularities in the plate sensitivity by appropriate calibration, and observe in (or reduce our observations to) the same zenith distance. Suppose we measure $S_{\lambda\lambda'}(l'/l)$ at two phases ψ_1 and ψ_2, and form their ratio. With the above precautions, $S_{\lambda\lambda'}$ can be regarded as a common multiplicative constant and cancels. Therefore

$$\begin{aligned}\left(\frac{l'}{l}\right)_1 : \left(\frac{l'}{l}\right)_2 &= Q \\ &= \left[\frac{1 + \Lambda(1 - \alpha)}{1 + \Lambda'(1 - \alpha')}\right]_1 \left[\frac{1 + \Lambda'(1 - \alpha')}{1 + \Lambda(1 - \alpha)}\right]_2;\end{aligned} \tag{5}$$

i.e., we obtain an equation expressing the ratio of the intensity gradients at different phases—a quantity the correct value of which can be deduced from spectrophotometric observation without much difficulty.

In order to introduce limb darkening explicitly we have to evaluate α. Provided that the law of darkening is of the form (2), it is easy to show that

$$\alpha^u = \frac{3(1 - u)}{3 - u}\alpha^0 + \frac{2u}{3 - u}\alpha^1, \tag{6}$$

where α^0 and α^1 are fractional losses of light of an "uniform" ($u = 0$) and completely darkened ($u = 1$) disc, respectively. As is well known, α^0 can be integrated directly in terms of circular functions, while for evaluating α^1 recourse to elliptical functions is inevitable. Both α^0 and α^1 have been extensively tabulated[8] as functions of the geometrical elements of the eclipse.

If we put

$$\frac{2u}{3 - u} = x \tag{7}$$

and abbreviate

$$\left.\begin{aligned} 1 - \alpha^0 &= f \\ \alpha^1 - \alpha^0 &= g \end{aligned}\right\} \tag{7.1}$$

equation (6) takes the form

$$\alpha^u = 1 - f + gx, \tag{6.1}$$

and equation (5) can be reduced to a standard form

$$Ax + Bx' + Cxx' = D, \tag{8}$$

the coefficients of which are:

$$\begin{aligned} A &= \Lambda\{g_2Q(1 + f_1\Lambda') - g_1(1 + f_2\Lambda')\}, \\ B &= \Lambda'\{g_1Q(1 + f_2\Lambda) - g_2(1 + f_1\Lambda)\}, \\ C &= \Lambda\Lambda'g_1g_2(1 - Q), \\ D &= Q(1 + f_2\Lambda)(1 + f_1\Lambda') - (1 + f_1\Lambda)(1 + f_2\Lambda'). \end{aligned}$$

Both f and g are purely geometrical functions of the phase angle, and Q is an observable quantity. The ratio of luminosities Λ also can, under certain conditions (*cf.* Section 6), be determined from the observations alone. Therefore the coefficients of equation (8) can be regarded as known, and the equation itself used for evaluating the x's and, with the aid of (7), the coefficients of limb darkening.

The procedure as developed so far is applicable no matter whether the eclipse in question is partial or total. It is easy to see that, in the case of partial eclipses, observations from three phases at least are required to yield a linear relation between x and x' (that is, to eliminate the quadratic term), and observations from six phases at least are necessary for obtaining the unknowns. This is of course in theory; in practice, many more observations would be required.

4.—If, however, the eclipse under consideration is total and spectrophotometric measurements extend also to the interval of totality, the problem admits of considerable simplification. During totality, obviously,

---

[8] Uniform tables, correct to four decimals, were computed by Wend (Diss. Leipzig, 1931) and Merrill (unpublished), while α^1 has been evaluated to the same number of digits by Ferrari (Wien Mitt., 1, 422, 1939) and, correctly to five decimals, by Zessewitch (*Bull. Astr. Inst. Leningrad,* No. 45, 1939).

$$(l'/l)_0 = L_2'/L_2;$$

if we express the intensity gradient in terms of $(l'/l)_0$ as unity,[9] f_0 and g_0 become zero and the coefficient of the quadratic term in (8) vanishes. The latter reduces to

$$A_0 x - B_0 x' = D_0 \tag{8.1}$$

where

$$\begin{aligned} A_0 &= g\Lambda \\ B_0 &= g\Lambda' Q \\ D_0 &= 1 + f\Lambda - Q(1 + f\Lambda'). \end{aligned}$$

We see that, in this case, observations in n phases, combined with those of totality, supply us with a system of n linear equations, a least squares solution of which yields the most probable values of the unknowns. Thus, in practice, the method can be applied with a reasonable expectation of success only when the minima studied are total.

Professor H. N. Russell has pointed out to the writer that, in the equations of condition (8.1) as they stand, the absolute terms D_0 are of very different weights. Since Q is determined photometrically, it is reasonable to assume that its error is proportional to itself, and the error of D_0 proportional to $Q(1 + f\Lambda')$. But, from (5) and (6.1),

$$Q = \frac{1 + (f - gx)\Lambda}{1 + (f - gx')\Lambda'} \, . \tag{5.1}$$

If we divide equation (8.1) by $(1 + f\Lambda)$, the error of its absolute term will be proportional to

$$\frac{1 - \dfrac{gx}{1 + f\Lambda}}{1 - \dfrac{gx'}{1 + f\Lambda'}} \, .$$

Since g is always a small quantity and $x < 1$, while f is always positive, we see that the weights of equations thus obtained will be approximately equal. We should therefore adopt, as our equations of condition,

$$A_{01} x - B_{01} x' = D_{01}, \tag{8.11}$$

---

[9] The maximum light could be used as the standard phase as well, but as $f_{\max} = 1$, the coefficients of Equation (8.1) would be more complicated.

where

$$\frac{A_0}{A_{01}} = \frac{B_0}{B_{01}} = \frac{D_0}{D_{01}} = 1 + f\Lambda.$$

This consideration applies only to errors arising from the inaccuracy of the underlying photometric measurements. But, intrinsically, equations (8.11) are still of very different weight. In order to illustrate the point, let us compute from (5.1) the theoretical Q's under the two limiting assumptions, namely $x = x' = 0$ (Q^0) and $x = x' = 1$ (Q^1), with any plausible set of geometrical elements. The reader may verify that the difference $Q^0 - Q^1$, zero at the moment of inner contact, increases rapidly with increasing ψ, but soon reaches a maximum and then begins to diminish again, until, in early partial stages, $(Q^0 - Q^1)$ becomes so small as to make any determination of the proper intermediate degree of darkening effectively impossible. Therefore equations corresponding to phases close to totality should be given more weight than those corresponding to early partial stages. It seems reasonable to put their weight proportional to the difference $(Q^0 - Q^1)$, which can be easily evaluated from (5.1) before equations (8) are solved.

The coefficients A_0 and B_0 are evidently mutually-dependent. The nature of the problem is such that the difference $(A_0 - B_0)$ depends strongly on the phase. Near totality, $(A_0 - B_0)$ is usually of the same order of magnitude as A_0 and B_0; but with increasing phase A_0 and B_0 approach asymptotically. This renders observations at early partial stages nearly useless for determining the individual values of x and x', but, as will be shown, they are still capable of providing us with a significant difference $(x - x')$. The writer is indebted to Professor Russell for the further suggestion that if, instead of x and x', new variables are introduced, defined as

$$\left.\begin{aligned} 2\bar{x} &= x + x' \\ \Delta x &= x - x' \end{aligned}\right\} \tag{9}$$

equation (8.11) then takes the form

$$A_{02}\bar{x} + B_{02}\Delta x = D_{02} \tag{8.12}$$

where

$$A_{02} = A_{01} - B_{01}; \qquad 2B_{02} = A_{01} + B_{01}; \qquad D_{02} = D_{01}.$$

The coefficients of $\bar{x}$ and Δx are practically always of the same sign, but the sum of A_0 and B_0 is, in every case, considerably greater numerically than their difference. Hence the nature of our problem is such that the difference between the coefficients of darkening in any two wave lengths can be obtained, with relatively far greater accuracy, from observations uniformly distributed along the whole light curve than can the individual coefficients.

5.—Our final consideration should concern errors possibly introduced by the inaccuracy of the geometrical elements. The latter enter into our procedure through f and g (*i.e.*, α^0 and α^1). We have seen in the preceding paragraph that the all-important phases (possessing greatest intrinsic weight) are those close to the inner contact. As is well-known, near totality α^0 and α^1 vary approximately as

$$\left.\begin{aligned} \alpha^0 &\doteq p^2 \\ \alpha^1 &\doteq 1 - (1 - p^2)^{\frac{3}{2}} \end{aligned}\right\} \qquad (10)$$

where

$$\left.\begin{aligned} p &= (\delta - r_2)/r_1 \\ \delta^2 &= \sin^2\psi \sin^2 i + \cos^2 i \end{aligned}\right\} \qquad (11)$$

r_1, r_2 being fractional radii of the primary and secondary component, respectively; δ, the separation of their centers (orbital radius taken as the unit of length); and i, the orbital inclination. Therefore, near totality,

$$\left.\begin{aligned} f &\doteq 1 - p^2 \\ g &\doteq (1 - p^2)(1 - \sqrt{1 - p^2}). \end{aligned}\right\} \qquad (7.11)$$

Differentiating, we obtain

$$\left.\begin{aligned} \Delta f &= -2p\left(\frac{\partial p}{\partial r_1}\Delta r_1 + \frac{\partial p}{\partial r_2}\Delta r_2 + \frac{\partial p}{\partial i}\Delta i\right) \\ \Delta g &= -2p\left(\frac{\partial p}{\partial r_1}\Delta r_1 + \frac{\partial p}{\partial r_2}\Delta r_2 + \frac{\partial p}{\partial i}\Delta i\right) \\ &\qquad\qquad \times\left(1 - \frac{3}{2}\sqrt{1 - p^2}\right) \end{aligned}\right\} \qquad (7.12)$$

where

$$\frac{\partial p}{\partial r_1} = \frac{p}{r_1}; \qquad \frac{\partial p}{\partial r_2} = -\frac{1}{r_1};$$
$$\frac{\partial p}{\partial i} = \frac{1}{r_1}\frac{\partial \delta}{\partial i}; \qquad \frac{\partial \delta}{\partial i} = -\frac{\sin i \cos i \cos^2\psi}{\delta}.$$

At the moment of inner contact ($p = -1$), the errors Δf and Δg are at maximum, namely,

$$\Delta f = \Delta g = -\frac{2}{r_1}\left(\Delta r_1 + \Delta r_2 + \frac{\partial \delta}{\partial i}\Delta i\right); \qquad (7.121)$$

but as $Q^0 = Q^1$, the weight of the corresponding equation of condition is zero. With diminishing eclipse, p increases from -1 to $+1$. But Δg simultaneously decreases until, at approximately $p = -0.745$, $(\partial g/\partial p)$ becomes zero and Δg vanishes, regardless of the errors in the geometrical elements.[10] The reader may verify that between $-0.80 < p < -0.70$ the differences $(Q^0 - Q^1)$ are usually at maximum, and the corresponding equations of condition are consequently of greatest weight. The fact that the coefficients A_0 and B_0 are, at these all-important phases, nearly independent of errors in the geometrical elements contributes greatly to the usefulness of the proposed method. With respect to the errors in f (occurring only in D_0), with increasing p, Δf decreases likewise; but, since f simultaneously increases, the quantity $\Delta f/f$ diminishes very rapidly. As long as f is small, the quantity $(1 + f\Lambda)$ is insensitive to f; if both f and Λ are large, $(1 + f\Lambda)$ may vary approximately as $+f\Lambda$, but in this case $\Delta f/f$ is likely to be so small that the effect of Δf will scarcely be troublesome in practical cases.

III. Application to Practical Cases. U Sagittae

6.—Application of the theory developed in the preceding section to practical cases requires not only that sufficiently precise measurements of Q be available, but also that the Λ's and the geometrical elements of the respective systems be independently known. The method proved to be practicable only for systems the deep minima of which are total (Section 4). Theoretically, the geometrical properties of such systems can be investigated without recourse to the partial branches of the light curve affected by a previously unknown amount of darkening. The radii ratio k and the ratio of luminosities Λ can be derived from the fractional losses of light in both minima, while r_1, r_2 and i can be evaluated from the observed angles of inner and outer contacts. This is in theory; in practice, secondary minima are frequently too shallow and the moments of outer contacts too ill-defined to make the procedure sufficiently accurate. Recourse to the partial branches of the light curves is then inevitable. The procedure is one by approximations. We adopt a certain plausible value for the degree of limb darkening, derive the corresponding geometrical elements, and with their aid compute the darkening coefficients from the observed Q's, in the manner described in Part II. If the resulting u differs from that adopted at the outset, the procedure is to be repeated until an agreement is established.

---

[10] The reason being that the errors of the geometrical elements in α^0 and α^1 are correlated and cancel.

7.—One of the most suitable eclipsing systems to which the above method can be applied is undoubtedly U Sagittae. Several light curves of this variable have been published,[11] the best of them being probably the photometric curve by Wendell.[12] Complete spectroscopic observations are available [13] which show that the bright component of U Sagittae, of spectral class B9n, becomes totally eclipsed in the deep minimum by a less luminous subgiant exhibiting gG2 characteristics; therefore, in agreement with the photometric solution, the B9-star must be the smaller one.

The light curve of U Sagittae is perfectly symmetrical with respect to conjunctions and exhibits no indication of an orbital eccentricity greater than, say, 0.01. Therefore, for the purpose of an orbital solution, the orbit can be regarded as circular. The "uniform" and "darkened" solutions, based on Wendell's light curve, have been carried out by Shapley,[14] and, assuming $u = 0.6$, by the writer. The resulting values are:

	U	$0.6\,D$	D
r_1	0.220	0.230	0.238
r_2	0.291	0.295	0.298
i	90°	90°	90°
Λ	11.7	11.4	11.0

Elements corresponding to any other intermediate degree of darkening can easily be obtained by interpolation. The eclipses result central irrespective of darkening. In addition to these values, the ratios of the luminosities Λ at $\lambda\lambda 6100$ and 4300 were recently determined by Walter;[15] we have

$\lambda 6100$	$\Lambda = 6.2$	Walter
5300	11.4	Wendell (0.6 D)
4300	25.0	Walter

If we plot $\log \Lambda$ against λ, the relation is found to be very approximately linear, and expressible as

---

[11] Wendell, *H. A.*, **69**, 82, 1909; Nijland, *A. N.*, **229**, 354, 1937; Gadomski, *Warsaw Repr.* 25, 1935; Walter, *Zs. f. Ap.*, **16**, 167, 1938.

[12] *Op. cit.* (*Cf.*, also, Shapley, *Princ. Contr.* 3, 162, 1915).

[13] Fowler, *Allegh. Publ.* 3, 11, 1916; Joy, *Ap. J.*, **71**, 336, 1930.

[14] *Princ. Contr.* 3, 1915.

[15] *Zs. f. Ap.*, **16**, 167, 1938.

$$\log \Lambda = 2.812 - 0.000329\,\lambda, \tag{12}$$

where λ is expressed in Angströms. With the aid of this formula, Λ can safely be interpolated for any wave length between λλ4000 and 6000, with an uncertainty probably not exceeding two or three per cent.

In the course of the above determination of the geometrical elements for U Sagittae, the form of its components was supposed to be spherical. In reality, however, both components are rotationally and tidally distorted and their equilibrium forms are ellipsoids. The spectroscopic mass-ratio of U Sagittae is $m_1/m_2 = 3.3$ (Joy). With the aid of the above radii we can surmise at once that the departure of the small and massive primary from a sphere is likely to be minute, while the secondary must be strongly distorted. Application of the equilibrium theory of distorted polytropes[16] shows that if the density condensations of both components are pronounced, and if $a > b > c$ denote their semi-axes, we have

$$b_1/a_1 = 0.996 \qquad b_2/a_2 = 0.89$$
$$c_1/a_1 = 0.991 \qquad c_2/a_2 = 0.84$$

The distortion of the primary is indeed so small that, within the scheme of our accuracy, we can consider it a sphere. The distortion of the secondary, however, is relatively large and its effects cannot be neglected. Its disc, projected on a plane perpendicular to the line of sight, will be an ellipse of eccentricity varying with the phase; at quadrature, the eccentricity of projection will be equal to that of the meridional plane of the secondary component, while, in conjunction, it will be equal to that of its diametral plane. This would simulate an increase of the radii ratio with increasing ψ;[17] or (as Russell[18] has recently pointed out), with the radii ratio kept constant, the effect simulates increased limb darkening. Thus, the effect, on the light curve, of the unequal form of both components can be very approximately accounted for if, in the course of the derivation of the geometrical elements, a proportionally higher degree of darkening is adopted.

---

[16] *Cf.* Chandrasekhar, *M. N.*, **93**, 462, 1933.

[17] This can be easily verified by approximating the elliptical boundary of the secondary by an oscillating circle. The radius of the latter diminishes as ψ increases. Thus, if $r_2 > r_1$, r_1/r_2 increases.

[18] *Ap. J.*, **90**, 641, 1939.

8.—The orbital period of U Sagittae is known with a high degree of accuracy. Since its discovery by Schwab in 1901, this variable has been under nearly constant scrutiny, and its period has seemed to be constant. It was not until recently that Joy (*op. cit.*) suggested the possibility of a periodic term. The latest published investigation of the period of U Sagittae is by Gadomski.[19] He has found that the minima observed from 1901 to the present time are at best satisfied by the following (heliocentric) elements:

$$\mathit{Min.} = \text{J.D. } 2417130.4171 + 3^d.3806184 \cdot E - 0^d.0092 \sin (0^\circ.16 \cdot E + 42^\circ). \quad \text{(I)}$$

Quite recently the period changes of U Sagittae have been studied also by Jacchia on the basis of Harvard plate material. The writer is indebted to Dr. Jacchia for communicating the results of his investigation in advance of publication. Jacchia finds the elements best fitting all the material available to be:

$$\mathit{Min.} = \text{J.D. } 2417130.4148 + 3^d.3806184 \cdot E - 0^d.0040 \sin (0^\circ.107 \cdot E + 83^\circ); \quad \text{(II)}$$

that is, he finds Gadomski's period correct, but that there is a small shift in the zero epoch ($0^d.0023$, or about 3 min.), and a smaller amplitude in the periodic term. It is estimated that the epochs predicted with the aid of the elements (II) will not deviate from reality (within the years 1900–1940, at least) by more than $0^d.001$, which corresponds to about 0°.1 in the phase angle.

As regards the physical significance of the periodic term, we may examine first the possibility that it could be invoked to demonstrate the presence of a third body. There exist two spectroscopic orbits of U Sagittae: by Fowler[13] (Allegheny, 1907–1911), and Joy[13] (Mount Wilson, 1926–1928), and the resulting velocities of the center of mass are closely concordant. In addition, during totality, there is no indication of another spectrum superposed on that of the G2 companion. Hence any hypothetical third body could have only a very small mass. Such a conclusion is consistent with the coefficients of the periodic term in (II) which, interpreted as being due to the light equation, would lead to a mass-function

$$\frac{m_3^3 \sin^3 i'}{(m + m_3)^2} = 0.17 \cdot 10^{-7} \odot ,$$

---

[19] *Warsaw Repr.* 25, 1935.

where m is the combined mass of the eclipsing system; m_3, the mass of a hypothetical third body; and i', the inclination of the third orbit to a plane perpendicular to the line of sight. We note at once that unless i' is exceedingly close to 0°—which is too improbable—the mass-ratio m_3/m must be very small. Inserting $m = 8.7 \odot$ (Joy), we should have, very approximately, $m_3 \sin i' = 0.022 \odot$. No star of such a small mass appears to be known at present; but since observational selection operates strongly against the discovery of companions of small masses, in visual as well as spectroscopic pairs, the possibility of their existence cannot be excluded.

An alternative explanation for the periodic term in the photometric elements may be that the apsidal line revolves in a slightly eccentric orbit. The secondary minimum is too shallow to permit any conclusions about the direction of its displacement, and thus provide conclusive proof as to the correctness of this hypothesis. But if the hypothesis were correct, the coefficient of the sine term as derived by Gadomski would correspond to an orbital eccentricity $e = 0.009$; and its period, to a ratio of the revolution of the apsidal line to the orbital revolution $U/P = 2250$; whereas, according to Jacchia, the corresponding quantities are: $e = 0.004$, $U/P = 3360$. The two spectroscopic orbits quoted above point to somewhat higher orbital eccentricities (0.03–0.06), but Luyten's[20] rediscussion has shown that these values are scarcely significant, and that a circular orbit would fit the spectroscopic observations just as well. It is, nevertheless, interesting to note that the longitudes of periastron as derived by Fowler and Joy are not inconsistent with the longitude of periastron and the rate of apsidal motion following from the oscillations of the orbital period.

Let us suppose the hypothesis of apsidal motion to be correct. Application of the Cowling-Sterne formula for the rate of advance of the apsidal line yields the following values of κ (that is, constant depending on the mean density condensation of the two components; for their definition, *cf.*, for instance, Sterne, *M. N.*, **99**, 451, 1939):

$$\kappa = 0.0043 \text{ (Gadomski)} \qquad \kappa = 0.0029 \text{ (Jacchia)}$$

Gadomski's value corresponds to a density condensation specified by the polytropic index $n = 3.75$, and Jacchia's value to $n = 3.87$. The latter is probably nearer to reality.

---

[20] *Ap. J.*, **84**, 85, 1936.

9.—In 1934, U Sagittae became an object of spectrophotometric study at the Victoria Observatory by R. O. Redman. A series of beautiful slit and slitless spectrograms in different partial phases as far as totality were secured with the aid of the 72-inch reflector, the subject of study being the center to limb variations of the absorption line intensities.[21] In 1939, Dr. Redman kindly sent the whole series of spectrograms to Dr. Shapley, who passed them on to the present writer for further study.

Redman's observations extend over five minima, the summary of which is given below. The first column of the tabulation gives the calendar date; the second column, the epoch; the third, the computed heliocentric minimum; the fourth, the correction for the light equation in the sun-earth orbit; and the last column, the computed geocentric minima (in G.C.T. throughout).

Date	*E*	Min. Hel. J.D. 2427000 +	L.E.	Min. Geo. J.D. 2427000 +
			d	
1934 Aug. 8	3114	657.6629	−0.0039	657.6590
14	3116	664.4241	−0.0037	664.4204
18	3117	667.8047	−0.0035	667.8012
Sept. 14	3125	694.8496	−0.0020	694.8476
24	3128	704.9915	−0.0014	704.9901

Table I contains a review of the spectroscopic material available. Successive columns give the number of the spectrogram in the Victoria records; the kind of camera employed (see explanations below the Table); the plate (Eastman 33 or 40); the Julian Date of the mid-exposure; the interval from the nearest minimum (in fractions of a day); the corresponding phase angle; the mean zenith distance during the exposure; the weight of the spectrogram, being a rough measure of its quality.

For more details concerning the instruments, development of plates, etc., the reader is referred to Redman's paper cited above.* Exposure times of the slitless spectrograms were short, generally not exceeding 10 minutes. For the slit spectrograms, however, exposures ran according to the brightness of the object and width

---

[21] *See* Redman, *M. N.*, **96**, 488, 1936.

* The writer is indebted to Dr. Redman for a personal communication giving further details concerning the calibration of the plates.

TABLE I

	Sp. No. Vict.	Instr.	Plate	Date of Mid-Exp. 2427000+	Int. from Min.	Phase Angle	Mean Zen. Dist.	Wt.
Totality:					d	°	°	
	23330	s	E 40	667.792	−0.009	1.0	32.7	½
	23420	1SS	E 40	694.860	+0.012	1.3	61.5	1
	23331	s	E 40	667.820	+0.019	2.0	37.0	½
Partial Phases:								
	23419	1SS	E 40	694.800	−0.048	5.1	48.3	1
	23332	s	E 40	667.850	+0.049	5.2	42.9	1
	23228	s	E 40	657.722	+0.063	6.7	31.8	1
	23418	1SS	E 40	694.763	−0.085	9.1	40.4	1
	23333	s	E 40	667.887	+0.086	9.2	51.1	1
	23230	s	E 40	657.749	+0.090	9.6	29.4	½
	23334	s	E 40	667.904	+0.106	11.3	55.8	1
	23417	1SS	E 40	694.739	−0.109	11.6	35.9	1
	23335	s	E 40	667.917	+0.116	12.4	68.0	1
	23231	s	E 33	657.801	+0.142	15.1	36.4	1
	23336	s	E 40	667.946	+0.145	15.4	65.0	½
	23232	s	E 33	657.820	+0.161	17.1	32.8	½
	23416	1M	E 40	694.681	−0.167	17.8	29.3	1
	23233	s	E 33	657.855	+0.196	20.9	38.5	½
	23475	1S	E 40	704.737	−0.253	26.9	40.4	1
Full Light:								
	23474	1S	E 40	704.701	−0.289	30.8	34.1	1
	23473	1S	E 40	704.659	−0.331	35.3	29.6	1
	23286	s	E 33	663.783	−0.637	67.8	36.0	1
	23285	s	E 33	663.771	−0.659	70.2	29.6	1
	23284	s	E 40	663.757	−0.663	70.6	29.3	1
	23283	s	E 40	663.741	−0.679	72.3	29.2	1
	23282	s	E 33	663.726	−0.694	73.9	29.7	1
	23281	s	E 33	663.711	−0.709	75.5	31.2	1
	23280	s	E 33	663.708	−0.712	75.8	31.6	1

s = slitless spectrogram; dispersion 136 A/mm at Hγ. 1SS = slit spectrogram; 92.3 A/mm at Hγ. 1S = slit spectrogram; 50.8 A/mm at Hγ. 1M = slit spectrogram; 30.1 A/mm at Hγ.

of the slit, from 26 minutes up to 100 minutes during totality. The plates were calibrated for photometric purposes. Microphotometric tracings were secured from all of the spectrograms with the aid of the Harvard Observatory's micro-densitometer;[22] some of the negatives in important phases were traced two or three times and reduced independently, in order to minimize accidental errors arising from irregularities in the plate grain. The writer gratefully acknowledges the assistance of Mr. Lawrence H. Aller who ran all the tracings.

As we have seen in Section 4, the determinateness of a solution giving the limb-darkening coefficients increases with increasing difference between the Λ's corresponding to the two different wave lengths in which measurements were carried out; in consequence, λ and λ' should be as widely separated as possible. In our present case we are limited in the ultra-violet by the faintness, during totality, of the refrangible end of a G2 spectrum, and in the yellow by a rather abrupt decrease of the plate sensitivity. Accordingly, the two wave lengths selected for measurement of the relative intensities were λλ4050 (in the neighborhood of the helium line λ4026, visible on some of the plates), and 4750. In the reduction of the measurements, differential extinction was corrected by means of the atmospheric transmission coefficients as published in Handbuch der Astrophysik, **2**, I, 206. The material was too scarce to enable one to draw any conclusions regarding possible systematic difference between the slit and slitless spectrograms; but, owing to the fact that very broad slits were used throughout, these differences are presumably small and escaped unnoticed.

---

[22] For its description, *see H. A.*, **105**, 99, 1937.

10.—Table II contains an analysis of the observations in partial phases, and illustrates numerically the procedure of computing the darkening coefficients. Successive columns give: (1) the phase angle; (2) the separation of the centers of the two components δ; (3) the quantity $p = (\delta - r_2)/r_1$; (4) the value of the fractional decrease in intensity α^0, the eclipsed disc assumed to be uniformly bright; (5) the value of α^1 corresponding to a completely darkened disc; (6) the measured ratio of the intensity gradient defined as $Q = (l'/l)_\psi : (l'/l)_{totality}$; and (7) the intrinsic weight of the phase (proportional to $(Q^0 - Q^1)$).

TABLE II

ψ	δ	p	α^0	α^1	Q	Wt.
5.1°	0.089	−0.895	0.962	0.985	1.21	13
5.2	0.091	−0.886	0.958	0.982	1.23	13
6.7	0.117	−0.773	0.895	0.940	1.40	9
9.1	0.158	−0.595	0.787	0.844	1.51	4
9.2	0.160	−0.586	0.781	0.838	1.51	4
9.6	0.167	−0.556	0.761	0.819	1.53	2
11.3	0.196	−0.430	0.681	0.735	1.56	1
11.6	0.201	−0.408	0.668	0.720	1.57	1
12.4	0.215	−0.347	0.629	0.677	1.59	1
15.1	0.261	−0.147	0.504	0.531	1.62	½
15.4	0.266	−0.125	0.491	0.516	1.62	½
17.1	0.294	−0.004	0.418	0.430	1.66	⅓
17.8	0.306	+0.048	0.388	0.395	1.68	⅓
20.9	0.357	+0.270	0.266	0.247	1.68	¼
26.7	0.452	+0.683	0.079	0.053	1.65	⅛

Computation of the coefficients for fifteen equations of condition of the type (8.11) or p (8.12) is then straightforward. The values of f and g depend only on α^0 and α^1 as tabulated above. The ratios of the luminosities Λ are obtained from (12); we find

$$\begin{aligned} \lambda 4050: &\quad \Lambda = 30.2 \\ \lambda 4750: &\quad \Lambda' = 17.8 \end{aligned}$$

at $\psi = 0^0$.[23] For the sake of numerical illustration of the method, the equations of condition are given explicitly below; in order to facilitate the calculations, equations (4–6) and (7–15) were grouped into one. We thus have:

No. of Observations	Equation	Wt.
(1)	$0.324\,x - 0.230\,x' = 0.059$	1.3
(2)	$0.320\,x - 0.231\,x' = 0.054$	1.3
(3)	$0.326\,x - 0.269\,x' = 0.039$	0.9
(4–6)	$0.226\,x - 0.201\,x' = 0.029$	1.0
(7–15)	$0.116\,x - 0.108\,x' = 0.013$	0.5

and the normal equations take the forms:

$$\begin{aligned} 0.0859\,x - 0.0658\,x' &= 0.0133 \\ 0.0658\,x - 0.0507\,x' &= 0.0101 \end{aligned}$$

Solving them, we obtain:

$$\begin{aligned} x &= 0.4 \pm 0.2 \text{ (p.e.)} \\ x' &= 0.3 \pm 0.2 \text{ (p.e.)} \end{aligned}$$

Hence

$$\begin{aligned} \lambda 4050: &\quad u = 0.5 \pm 0.2 \text{ (p.e.)} \\ \lambda 4750: &\quad u' = 0.4 \pm 0.2 \text{ (p.e.)} \end{aligned}$$

These values were computed with the geometrical elements corresponding to 0.6 D. Owing to the distortion of the secondary component (which gives rise to an effect simulating increased limb darkening), these elements in reality correspond to a true degree of darkening smaller by 0.2–0.3 (*cf.* Russell, *op. cit.*, p. 670). Hence the agreement between the assumed and resulting darkening is good and no further approximations are necessary.

A glance at the above equations of condition shows, however, that as the eclipse diminishes the coefficients of x and x' approach rapidly a common limit—a fact which adversely affects the determinateness of the solution and decreases its weight. Yet, in Section 4 we have seen that the difference Δx can be determined under existing conditions more accurately than the individual x's. Let us, therefore, introduce $\bar{x}$ and Δx as unknowns and use (8.12) as our equations of condition, which then take the form:

---

[23] To the values of Λ has still to be applied a small correction due to the reflexion effect. In systems with large Λ, such as U Sagittae, the amount of the primary's radiation reflected from the secondary may represent a not insignificant contribution to the proper luminosity of the latter; in fact, the luminosity of the secondary will be increased by a small factor which can be theoretically evaluated (*cf.* the writer's forthcoming paper in the *Annals of the New York Academy of Sciences*, Vol. **41**; *H.R.* 211). As a result, the "effective" value of Λ will slightly decrease with increasing phase.

No. of Observations	Equation	Wt.
(1)	$0.094\,\bar{x} + 0.277\,\Delta x = 0.059$	1.3
(2)	$0.089\,\bar{x} + 0.276\,\Delta x = 0.054$	1.3
(3)	$0.057\,\bar{x} + 0.298\,\Delta x = 0.039$	0.9
(4–6)	$0.025\,\bar{x} + 0.214\,\Delta x = 0.029$	1.0
(7–15)	$0.008\,\bar{x} + 0.112\,\Delta x = 0.013$	0.5

As is evident, the coefficient of Δx exceeds that of $\bar{x}$ in every case, and their disparity increases rapidly as the eclipse diminishes. The normal equations then become:

$$\begin{aligned} 0.00518\,\bar{x} + 0.0177\,\Delta x &= 0.00318 \\ 0.0177\quad \bar{x} + 0.0672\,\Delta x &= 0.0115 \end{aligned}$$

and their solution is

$$\begin{aligned} \bar{x} &= 0.31 \pm 0.21 \text{ (p.e.)} \\ \Delta x &= 0.08 \pm 0.06 \text{ (p.e.)}. \end{aligned}$$

The latter corresponds to

$$u - u' = 0.09 \pm 0.06 \text{ (p.e.)}.$$

The probable error of a single color-index difference (*i.e.*, 2.5 log Q) results in the neighborhood of $0^m.01$.* The weights of both sets of solutions are:

$$\begin{aligned} \sqrt{p_x} &= 0.023 \qquad & \sqrt{p_x} &= 0.022 \\ \sqrt{p_{x'}} &= 0.018 \qquad & \sqrt{p_{\Delta x}} &= 0.086. \end{aligned}$$

Thus, the present example does show that in favorable cases, such as U Sagittae, a series of moderately accurate color indices can yield a relatively accurate difference between the darkening coefficients in the two particular wave lengths.

IV. APPLICATION TO U CEPHEI

11.—Another well-known system to which the present method can readily be applied is U Cephei. Geometrically it resembles very much that of U Sagittae.[24] It has already been the subject of a similar study by Rosenberg [7] who, however, did not investigate adequately the nature of the problem and contented himself with a semi-graphical procedure. His results for the darkening coefficients of the A0 component of U Cephei (λ4015, $u = 0.20$; λ6165, $u = 0.05$) are much smaller than those we have just found for U Sagittae (spectrum B9). However, a glance at Rosenberg's graphical representation of the residuals, from the general trend of the color-index variation during eclipse (*op. cit.*, p. 75, Fig. 3), shows that whereas the "uniform" solution exhibits residuals increasing in magnitude in the proper direction, in the "darkened" solution no significant residuals seem to remain. The latter solution should consequently be preferred, but is difficult to reconcile with coefficients as small as $0.05 < u < 0.20$, as postulated finally by Rosenberg (*op. cit.*, p. 81). Therefore, a re-discussion of Rosenberg's data by the method applied to U Sagittae appears very desirable. Unfortunately, Rosenberg did not publish his original observations of the color indices, but only a general trend of the color-index variation, presumably obtained by smoothing the individual observations; nor did he give the reader any idea as to the probable errors of any of his results. An application of the least-squares method to such material may seem contrary to the principles of the theory of errors. Nevertheless, provided that the smoothing was performed correctly, the procedure is still capable of yielding correct values for the unknowns, but is unable to give their true probable errors.

---

* The reader may recall that, in order to determine, with similar precision, the coefficient of darkening simultaneously with the geometrical elements, from a light curve observed in only one wave length, observations at least ten times as accurate are necessary.

The geometrical elements of U Cephei were, under the U and D hypotheses, derived by Rosenberg (*op. cit.*, p. 78, Tab. VII) from his light curves, and an additional solution corresponding to 0.6 D was carried out by the writer. We have:

	U	0.6 D	D
r_1	0.203	0.214	0.222
r_2	0.324	0.322	0.320
i	90°	90°	90°

λ 4015: $\Lambda = 22.76$
λ 6165: $\Lambda' = 4.97$

---

[24] Among the very numerous investigations of U Cephei (*cf. Geschichte und Literatur*, 1, 27; and *Ergänzungsband*, 1, 264), an exhaustive study by Dugan, *Princ. Contr.* 5, 1920, is probably the best.

Table III illustrates the procedure numerically. The headings of the columns (analogous to those of Table II) are self-explanatory. The values of Q were computed from Rosenberg's color indices (*op. cit.*, p. 74, Tab. V), with the use of the relation

$$\log Q = \tfrac{2}{5}\left[(C.I.)_{totality} - (C.I.)_{\psi}\right].$$

The observation at $0^d.040$ ($\psi = 5^\circ.8$; that is, approximately the moment of inner contact) was excluded, because it is questionable whether or not the eclipsed star begins at that moment to participate in the combined light of the system.

TABLE III

ψ	δ	p	α^0	α^1	Q	Wt.
7.2°	0.126	−0.915	0.976	0.991	1.33	2.3
8.7	0.151	−0.798	0.921	0.957	1.74	3.8
10.1	0.175	−0.686	0.858	0.906	2.20	3.4
11.6	0.202	−0.560	0.780	0.834	2.69	2.5
13.0	0.225	−0.453	0.712	0.764	3.03	1.8
14.4	0.249	−0.340	0.640	0.685	3.29	1.1
15.9	0.273	−0.228	0.570	0.604	3.48	0.8
17.3	0.298	−0.111	0.497	0.517	3.61	0.3

The weights of the observations for $\psi > 18^\circ$ (p positive) are so small as to make these phases virtually irrelevant for the determination of limb darkening.

The coefficients of the equations of condition of the type (8.11) or (8.12) can now be easily evaluated and are given in the following tabulation:

ψ	A_{01}	B_{01}	D_{01}	A_{02}	B_{02}
7.2°	0.221	0.064	0.036	0.057	0.143
8.7	0.293	0.111	0.134	0.182	0.202
10.1	0.258	0.124	0.113	0.134	0.191
11.6	0.205	0.120	0.062	0.085	0.163
13.0	0.157	0.104	0.025	0.053	0.131
14.4	0.112	0.080	0.001	0.032	0.096
15.9	0.073	0.056	−0.017	0.017	0.065
17.3	0.037	0.029	−0.015	0.008	0.033

The coefficients of the normal equations of the type (8.12) then are:

$[A_{02}A_{02}] = 0.0137$ $\qquad$ $[B_{02}B_{02}] = 0.0255$

$[A_{02}B_{02}] = 0.0186$ $\qquad$ $[B_{02}D_{02}] = 0.0142$

$[A_{02}D_{02}] = 0.0105$

Solving for the unknowns, we find

$$\bar{x} = 1.0, \qquad \Delta x = -0.15,$$

which are contradictory, since x' would have to exceed unity. However, it is easy to show that these results are of little significance. Their weights are:

$$\sqrt{p_x} = 0.013, \qquad \sqrt{p_{\Delta x}} = 0.017,$$

but

$$\sqrt{\frac{[wvv]}{[w](n-2)}} = 0.032,$$

that is, the residuals for U Cephei are about four times as large as the residuals for U Sagittae. The probable error of $\bar{x}$ is then ±1.7, and that of Δx, ±1.3, making both of them of no significance whatever. Rosenberg's material thus proves to be inconclusive as to the amount of limb darkening of the principal component of U Cephei or its variation with wave length. Deviations of the individual color indices from the general trend, interpreted by Rosenberg as quantitative indications of limb darkening (principally at $d = \pm 0^d.050$), are not systematic enough to give the solution sufficient weight.

V. General Discussion

12.—From the theoretical point of view, the ultimate aim of determinations of limb darkening coefficients is to investigate the extent of the variation of $\kappa_\lambda/\bar{\kappa}$ with the temperature or wave length. As regards the sun, it has been pointed out by Milne[25] and Lindblad [26] that the observations are best brought into agreement with theory on the assumption that κ_λ is practically independent of wave length and equal to $\bar{\kappa}$. The agreement becomes nearly complete if the blanketing effect of the reversing layer is taken into account.[27] Large variations of κ_λ with λ, as required by recent theories of the absorption coefficients, are definitely disproved in case of the sun; or, as was admitted by Pannekoek, "solar phenomena would correspond to those computed for an effective temperature of over 7000°." [28]

For YZ Cassiopeiae the darkening coefficient has been determined by Kron, so far for only one wave length (λ4500). Kron assumed the effective temperature of its primary component (spectrum A3) to be 10500°. According to Pannekoek, one should expect $u = 0.65$, while, if we assume a constant absorption coefficient, $u = 0.53$. The observed value is $u = 0.49 \pm 0.04$, favoring undoubtedly the latter alternative. It may be pointed out that an effective temperature of 10500° for an A3 star is probably somewhat overestimated; Kuiper[29] recently gave 9400°. With the latter temperature, the agreement between the observed and theoretical values, corresponding to a constant absorption coefficient, becomes slightly less evident, but is still far from confronting us with a serious discrepancy.

---

[25] *M. N.*, **81**, 375, 1921.

[26] *Upsala Univ. Arsskr.* 1, 1920.

[27] *Cf.* Milne, *Observatory*, **51**, 88, 1928.

[28] *M. N.*, **95**, 734, 1935.

The present results for U Sagittae seem to point to a similar conclusion. The darkening coefficient of the B9 component, within the interval $4050 < \lambda < 4750$, is found to be nearly constant, or to increase but slowly with increasing frequency of light. According to Pannekoek, u should increase from 0.53 at $\lambda 4700$ to 0.67 for $\lambda 4050$; if the absorption coefficient is constant, u should be expected to increase within the same interval from 0.55 to 0.57. The probable errors of the present results are rather large and do not permit us to distinguish between these alternatives unquestionably; but again the latter alternative is apparently more probable.

Thus, so far as the present observational evidence goes, the agreement between the observed darkening coefficients and those predicted under the assumption of a constant continuous absorption coefficient seems significant. The absorption coefficient need not, however, be strictly independent of wave length; there is no theoretical reason why it should be, and observational evidence is still far from being ideally accurate. But it seems evident that such large variations of κ_λ with λ, as postulated, for instance, by Pannekoek, are scarcely admissible.

In conclusion, the writer takes pleasure in expressing his thanks to Dr. R. O. Redman for lending his spectrograms for the foregoing study of U Sagittae, and to the late Dr. W. E. Harper, former Director of the Dominion Astrophysical Observatory, for authorizing their use. The writer is further deeply indebted to Professor Henry Norris Russell for his stimulating interest and valuable suggestions, to which the present paper owes a great deal. He would like also to express his appreciation to Mr. Lawrence H. Aller for his assistance in making the micro-densitometric tracings.

---

[29] *Ap. J.*, **88**, 429, 1938.

N10

THE PRODUCTION OF NEUTRONS BY THE COSMIC RADIATION

S. A. KORFF

Bartol Research Foundation of The Franklin Institute, Swarthmore, Pa.

Abstract

The present state of knowledge about cosmic-ray neutrons is reviewed. The altitude dependence of the rate of production is investigated, employing neutron counters and the radio-balloon technique. A rate of increase with elevation is found which is equal to that of the soft component. Other experiments are described in which the coincidence between neutrons and showers is observed at two elevations, by observing the simultaneous discharge of neutron and shower counters. The calibrations and controls on the experiments are described. It is further shown that a cross section of less than 10^{-25} cm sq. will suffice to account for the observed number of neutrons if they are produced by photons. In summary, the altitude dependence, the observed coincidences and the cross section considerations all point to a production of the neutrons by the soft component presumably by a nuclear photodisintegration process.

Introduction

Neutrons constitute one of the important components of the cosmic radiation. Experiments have shown that they are approximately equal numerically to the charged particles in the radiation. Studies have been made by Rumbaugh and Locher[1] who sent photographic plates up to an elevation of 72,000 feet at the time of the National Geographic-U. S. Army Air Corps flight of the balloon Explorer II, and also by Schopper[2] and others, using similar technique. In these experiments a rapid increase of the neutron intensity with elevation was observed, but a quantitative interpretation of the photographic plate data in terms of an absolute rate of neutron-production is difficult, owing to our scant knowledge of the range-energy relationship of various nuclear particles in gelatine.

The neutron intensity was also measured by Fünfer[3] who employed a boron-lined ionization chamber and linear amplifier near sealevel and at the top of the Zugspitze. Von Halban, Kowarski and Magat[4] conducted an airplane flight over Paris, up to an elevation of 30,000 feet (3 meters of water equivalent below the top of the atmosphere), in which they measured the neutron intensity by observing the induced radioactivity in bromine. At the Bartol Foundation, the neutrons at sealevel have been measured by C. G. and D. D. Montgomery,[5] and at high elevations by the radio balloon technique by Korff.[6] The results of these measurements have been analyzed by Bethe, Korff, and Placzek,[7] who have discussed the history of neutrons in the atmosphere, and who have shown what energy distributions might be expected. Whereas all of these observations agree on the numerical importance of the neutron component of the radiation, very little is known at the present time regarding the process by which these neutrons were produced. The experiments to be described were made for the purpose of throwing further light on the mechanisms of production.

---

[1] L. H. Rumbaugh and G. L. Locher, *Phys. Rev.*, **44**, 855 (1936).

[2] E. and L. Schopper, *Phys. Zeits.*, **40**, 22 (1939).

[3] E. Fünfer, *Zeits. f. Phys.*, **111**, 351 (1938).

[4] H. v. Halban, L. Kowarski and M. Magat, *Comptes Rendus*, **208**, 572 (1939).

The neutrons are presumably produced in the atmosphere by the high energy cosmic radiation. Neutrons are not believed to be primary particles because of their theoretical instability. Moreover, it would be difficult to account for the observed altitude dependence and energy distribution on the assumption that these neutrons are originated outside the atmosphere. We may, therefore, examine the possible production processes and consider the experimental evidence which may disclose the nature of such processes.

There are two chief lines of evidence which provide information regarding the production mechanisms. These are, (*a*) the number of neutrons as a function of elevation, and (*b*) the observed coincidences between neutrons and showers. Some new experimental data on each of these two are now presented.

APPARATUS

1. *Balloon Apparatus.*—The instruments used in investigating the altitude dependence of the neutron intensity consist of a neutron counter, an attached amplifier and pulse prolonging circuit, a radiobarograph, and a radio transmitter. The instruments are carried aloft by free balloons and the neutron intensity as well as the atmospheric pressure is automatically transmitted

---

[5] C. G. and D. D. Montgomery, *Phys. Rev.*, **56**, 10 (1939).

[6] S. A. Korff, *Revs. Mod. Phys.*, **11**, 211 (1939).

[7] H. A. Bethe, S. A. Korff, and G. Placzek, *Phys. Rev.*, **57**, 573 (1940).

by short wave radio to the receiving station on the ground. This technique has been previously described in detail.[6,8]

The counters used in this work are filled with boron trifluoride gas. A neutron captured by one of the boron nuclei will produce a disintegration and the resulting alpha particle may be counted by the proportional counting technique. Such a counter will have a background due to the alpha particles produced by natural radioactive contamination of the walls of the cylinder. This background may be determined by observing the counting rate in the absence of neutrons. Further calibrations may be obtained by operating such counters in the presence of a known source of neutrons, due allowance being made for the geometry. The corrections to which these counters are subject have been discussed elsewhere.[7]

2. *Coincidence Measurements.*—If there is a connection between the process of neutron production and that of the production of showers, then it should be possible to observe coincident discharges of neutron counters and shower detecting units. In order to measure the coincidences between neutrons and showers, a neutron counter and a shower recording tray were so connected that coincident discharge was necessary in order to record. The neutron counter was one filled with boron trifluoride gas, of the type discussed above, 7 cm in diameter and 42 cm long. The pulses from this counter were fed into the conventional type of counter amplifier circuit. A shower tray, consisting of 25 Geiger counters arranged in parallel according to the method described by Ramsey,[9] were connected to a circuit so adjusted that at least two of these counters had to discharge in order to record as a count.

The experimental arrangement is shown in Fig. 1, in which the neutron counter will be seen. Under the neutron counter is a slab of lead, to act as shower producing material. The neutron counter [10] is placed above the lead in order to minimize the effect of large showers produced in the lead, which might cause the counter to discharge. A neutron produced in the lead will have a sufficient probability of being initially produced in an upward direction, whereas practically all of the showers produced in the lead will be directed downward. The neutron counter is, in

---

[8] S. A. Korff, *Rev. Sci. Inst.*, **9**, 256 (1938). S. A. Korff and T. H. Johnson, *Rev. Sci. Inst.*, **10**, 82 (1939).

[9] W. E. Ramsey, *Phys. Rev.*, **57**, 1061 (A), (1940).

[10] S. A. Korff and W. E. Danforth, *Phys. Rev.*, **55**, 980 (1939).

addition, surrounded by paraffin for the purpose of slowing down the neutrons and thus increasing detection efficiency, and also to scatter neutrons into the counter.

Below the slab of lead is a shower tray, and below this is additional paraffin which may scatter backward any neutrons ejected downward from the lead. The neutron counter will not respond to the passage through it of the ordinary particles of the cosmic radiation, or of showers of moderate size, the limits being discussed below, nor will the shower tray detect neutrons. A coincidence may be taken to indicate that at least two ionizing particles have passed through the shower tray at the same time that one neutron was recorded by the neutron counter, the simultaneous ionizing particles being presumptive evidence that a shower was created in nearby matter.

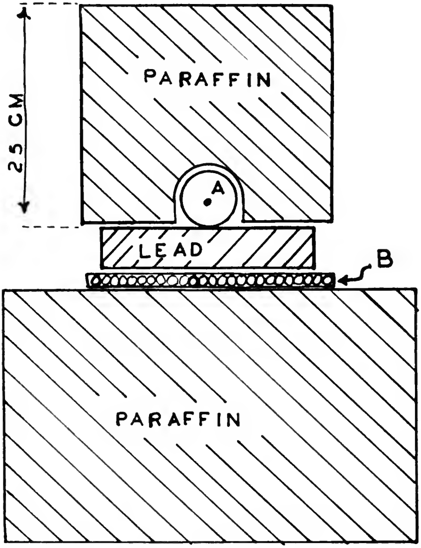


FIG. 1. Geometrical arrangement used in studying neutron-shower coincidences. *A* is a neutron counter, *B* is a shower tray. Counting rates of each are determined separately, and rate of coincidences is observed. Showers may be generated in the lead block. Neutrons are slowed and scattered in the paraffin.

The circuit used is shown in Fig. 2. It will be seen that this is a normal coincidence circuit, with a resolving time of about 10^{-4} second. This time was selected from a consideration of the mean lifetime of neutrons in paraffin in order to increase the efficiency of the unit by allowing neutrons to be counted after they had been slowed down in the paraffin.

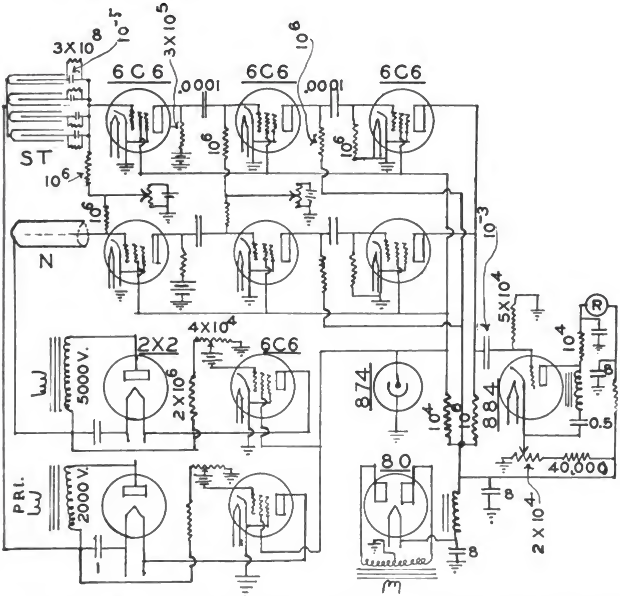


FIG. 2. Diagram of circuit used to study neutron-shower coincidences. Resistances are in ohms, capacities in microfarads. Pulses from neutron counter (*N*) and shower tray (*ST*) are fed through standard three-stage amplifier and coincidence circuit, and operate a thyratron recorder (*R*). Stabilized voltage source below.

OBSERVATIONS

The observations on the neutron intensity made with the aid of balloon flights are presented in Table 1. The observed number of counts per second (n) of the counters in the several experiments, the counter volume (V), and the pressure (p) in atmospheres of the BF_3 in the counters are listed. The rate of production (q) per gram of air per second of the neutrons by the cosmic radiation is determined by (n) through the equations discussed below. In these balloon flights the neutron counter is in the free atmosphere and no paraffin is used.

TABLE 1

OBSERVED COUNTING RATE OF NEUTRON COUNTER AS A FUNCTION OF ELEVATION AND VALUES OF q, THE RATE OF PRODUCTION OF NEUTRONS IN AIR BY THE COSMIC RADIATION, FROM BALLOON FLIGHT DATA

Flight number	Elevation, meters	Atmospheric depth, water equivalent	n observed neutron-counts per sec.	V counter volume, cc	p BF_3 atm.	q rate of production per gram per sec.
24	4200 m	6 m	1/120	147	0.158	6.5×10^{-4}
24	7200	4	1/40	147	0.158	1.9×10^{-3}
21	7200	4	1/60	240	0.075	1.7×10^{-3}

The observed rates of production of neutrons in paraffin (Q) on Mount Washington and near sealevel are shown in Table 2. The value of Q is for paraffin, and is calculated from eq. 1 below.

TABLE 2

THE OBSERVED RATES OF PRODUCTION OF NEUTRONS IN PARAFFIN AT TWO ELEVATIONS

Note that Q should not be compared with q, without considering transition effects.

Elevation, meters	Depth	n	V	p	Q
Bartol, 100 m	10.4	0.017	1450	0.5	3.3×10^{-5}
Mt. Washington Observatory, 1925 m	8.0	0.1	1450	0.5	1.9×10^{-4}

In this case, (n) is the difference between the counting rates of a neutron counter, with and without a cadmium shield, and is therefore a measure of the slow neutrons in the paraffin.

Finally, the observed rates of coincident discharge of the neutron and shower counters for two elevations are presented, as

TABLE 3

COUNTING RATES OBSERVED IN NEUTRON-SHOWER COINCIDENCE EXPERIMENTS AT TWO ELEVATIONS, AS WELL AS CONTROL EXPERIMENTS

Elevation	Depth, meters of water	Counting rate, counts per second		
		Neutron counter alone (C_n)	Shower tray alone (C_s)	Neutron-shower coincidences (C_c)
Bartol, 100 m	10.4	0.017	0.14	3.8×10^{-5}
Mt. Washington	8.0	0.1	0.74	2.5×10^{-4}
Bartol: Controls: Cadmium shield on neutron counter				0 in 22½ hours
BF$_3$ counter voltage lowered				0 in 20 hours

well as the individual counting rates of the neutron counter and shower tray, in Table 3.

Discussion

A. Altitude Dependence

The results of the measurements of the rate of neutron production at various altitudes are plotted in Fig. 3, as a function of atmospheric depth in meters of water equivalent. On the same diagram the altitude dependences of (*a*) the total cosmic ray intensity, from the data at corresponding geomagnetic latitudes

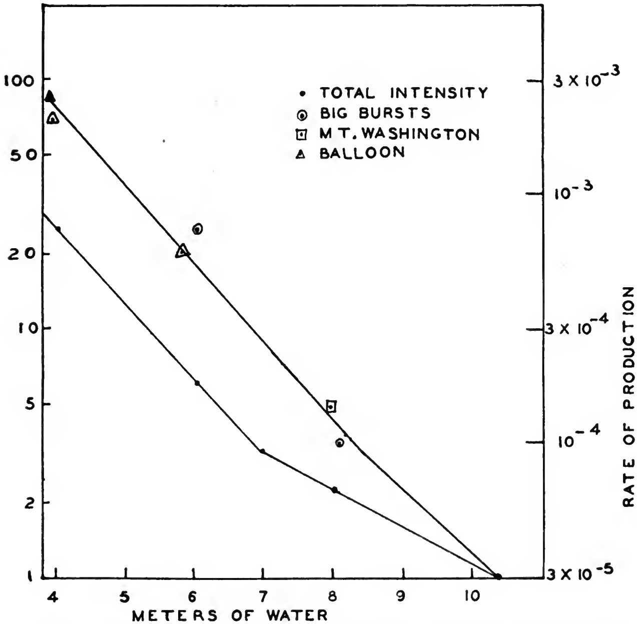


FIG. 3. Neutron intensity as a function of altitude. Left-hand ordinates, relative intensities, sealevel being unity. Right hand ordinates, rate of production of neutrons, q, per gram per second by the cosmic radiation. Lower curve, total cosmic ray intensity. Circles, big bursts. Parallelism of altitude dependence of neutrons and soft component indicated.

obtained by Bowen, Millikan and Neher,[11] and (*b*) the large bursts, from the observations of Montgomery and Montgomery,[12] are also plotted. It will be noted that the rate of increase of the neutron-production with altitude is roughly the same as that of the big bursts, while it is faster at low elevations than the rate of increase of the total cosmic ray intensity. At higher elevations, where the soft component produces the bulk of the total ionization, the total intensity and the neutron production increase at the same rate.

The left hand ordinates are relative intensities applicable to all the various quantities discussed, and plotted on a logarithmic scale. All altitude dependences are expressed in terms of a common point, the intensity of each at sealevel being taken as unity. The right hand ordinates are applicable only to the neutron intensities and are the values of q, the rate of production of neutrons per gram per second in the atmosphere. They are computed from eq. 12 of the analysis made by Bethe, Korff and Placzek.[7] In this analysis it was shown that the counting rate of (n) counts per second of a counter in air is related to (q) through the equation:

$$q = (n/V)(\sigma_A/\sigma_D)(780/p), \tag{1}$$

where V is the volume of the counter, p the pressure in atm. of BF_3 therein, σ_D the capture cross section of BF_3 for neutrons and σ_A that for air. Taking account of the isotope ratio in commercial BF_3, we may take σ_D as approximately 550×10^{-24} sq. cm. If we take σ_A as 1.5×10^{-24}, then the observed (n) and V and p in Table 1 determine the values of q there listed.

We may next consider the neutron observations in Table 2. In this case, a large amount of paraffin surrounds the counter, with the result that few neutrons from the outside will reach the counter, but instead the counter will measure those produced in the paraffin. Since carbon and nitrogen nuclei are quite similar, we may expect that the rate of production of neutrons by the high energy cosmic radiation will not be very different in the two substances, and we may use eq. 1, substituting for σ_A the capture cross section for paraffin. If we take this to be 10^{-24} sq. cm, then the values of Q for paraffin will be those in Table 2. We must recall, however, that while the paraffin surrounding the counter is thick for neutrons, it is not thick for the cosmic radiation. Whatever component produces neutrons may be expected to show a transition effect, and the values for q and for Q are therefore not necessarily comparable.

---

[11] I. S. Bowen, R. A. Millikan and H. V. Neher, *Phys. Rev.*, **53**, 855 (1938).

[12] C. G. and D. D. Montgomery, *Phys. Rev.*, **47**, 429 (1935).

B. Coincidence Measurements

We will now discuss the interpretation of the neutron-shower coincidence experiment. Let us designate by C_s the counting rate of the shower tray when operating by itself, by C_n that of the neutron counter when operating by itself, and by C_c the counting rate of the arrangement in coincidence. Further, let E_s be the efficiency of the shower tray in detecting showers, and E_n that of the neutron counter in counting neutrons, so that the total number of showers will be C_s/E_s, and the total number of neutrons C_n/E_n. Designate by Z the total number of events in which both a shower and a neutron are produced, so that $Z = C_c/E_nE_s$. We may now inquire what fraction of the total number of showers S are showers in which a neutron is produced. This fraction will be:

$$Z/S = C_c/E_nC_s. \tag{2}$$

We may also inquire what fraction of all the neutrons (n) are produced in processes which are accompanied by showers. This will be:

$$Z/n = C_c/E_sC_n. \tag{3}$$

It will be recalled that the efficiency of such a neutron counter in counting slow neutrons will be given by:

$$E_n = Mp\sigma_D L, \tag{4}$$

where M is the loschmidt number, L is the length of the average path for neutrons passing through the counter, and p is the pressure of BF_3 in atm. in the counter. In this case, p was 0.5 and L was 7 cm. The value of σ_D depends on the neutron energies according to the $1/v$ law, and if the neutrons are produced near the counter, as for example in the lead, those passing through the counter will have more than thermal energies. If we suppose that the average neutron has ten times thermal velocity, E_n is 5×10^{-3}. If all neutrons are thermal, E_n would be 5×10^{-2}.

Using the lower value for E_n and the observed C_c and C_s, we get for Z/S about 6×10^{-2}. This would suggest that roughly 6 in every 100 showers have a neutron-producing event associated. If all neutrons were thermal, Z/S would be 6×10^{-3}, a definite lower limit for this quantity being thus established.

Similarly we may proceed to ascertain whether every neutron-producing event is accompanied by a shower. Our knowledge of E_s is somewhat uncertain, for it is impossible to allow for showers, one ray of which passes through the tray, or others in which several rays may pass through a single counter. Also, neutrons produced in showers produced below the tray might be detected, while the showers would not. In the limits, if E_s is 2.7×10^{-3}, then Z/n is unity and each neutron-producing event is accompanied by a shower; while if E_s is taken at 0.1, then Z/n will be 2.7×10^{-2}.

We may also compare the rate of production of neutrons with that of photons in the lead slab. The total number of electrons below a lead slab has been measured by Swann and Ramsey.[13] We can assume that the number of photons is equal to that of the electrons. The observations cited include all electrons, both those directly striking the lead, and those produced as knock-ons by mesotrons in the lead. They find 8×10^{-2} electrons per cm. sq. per minute below a piece of lead 5 cm thick. Since a 5 cm column contains about 10 unit lengths in radiation theory, 0.8 photons are produced per column per minute, or about 2.4×10^{-4} per gram per second.

We may roughly estimate the rate of production of neutrons in the lead. The counter surrounds about 0.1 of the total lead area, and hence the probability that a neutron ejected in a random direction will pass through the counter and be detected is 5×10^{-4}. Then in 2×10^4 grams of lead, $C_c/5 \times 10^{-4} \times 2 \times 10^4$ gives 3.8×10^{-6} neutrons per gram per second. Hence the ratio of the rate of production of neutrons to photons is 1.6×10^{-2}, if one neutron is produced per event. This figure lies within the limits for Z/S above.

It must again be emphasized that these estimates are necessarily somewhat uncertain, since it is impossible to make exact allowance for the different way in which geometry will affect neutrons and the showers, and since we cannot compute E_n or E_s accurately.

It will be recalled that heavy nuclear particles were observed in Anderson's cloud-chamber photographs,[14] made on the top of

---

[13] W. F. G. Swann and W. E. Ramsey, *Phys. Rev.*, **57**, 749 (1940) and **58**, 477 (1940).
[14] C. D. Anderson and S. H. Neddermeyer, *Phys. Rev.*, **50**, 263 (1936).

Pike's Peak, occurring in 113 out of 9188 pictures. Since his cloud-chamber was counter-controlled, about one per cent of the ionizing events actuating his apparatus were accompanied by the generation of nuclear particles. This differs by less than an order of magnitude from the values of Z/S discussed above. Thus Anderson's protons are produced at roughly the same rate as neutrons, and neutrons would, in general, not be seen in his photographs because of the small recoil-cross section. Presumably the protons and neutrons have a common origin in a nuclear disintegration process.

In an attempt to find out whether neutrons were principally associated with large or with small showers, the shower tray was so adjusted that at least four counters had to discharge in order to record, instead of two as before. With this arrangement, the counting rate was only about 50 per cent less, in coincidence, than its previous value, while the C_s for four ray showers was about one fourth of C_s for two ray showers. Hence the neutrons appear to be associated chiefly with the large showers.

C. Factors Affecting the Experiment

It is necessary to consider the effects due to gamma rays. In all cases, following the customary procedures in proportional counting technique,[15] the neutron counter was so adjusted that it did not count electrons. That it did not, could be easily ascertained by bringing some radium near it. It will be recalled that the gamma rays from radium will produce secondary electrons by various processes in the counter. The largest amount of ionization which an electron can produce is that which it creates when it ends its range along the length of the counter. In most of the counters used, this corresponded to about 50 Kev, and hence the counter-voltage was so selected that a 50 Kev ionizing event was not large enough to record.

In the case of the balloon flights, a calibration was made in the following way. The neutron counting rate was determined when a 10 milligram radium-beryllium neutron source was in a lead block 8.5 cm in radius and located 40 cm from the counter. The counter was surrounded by approximately 6 cm of paraffin arranged with irregular geometry. The neutron counting rate was then observed. Next a 2 milligram radium source was brought within 1 meter of the apparatus and no detectable change was observed in the neutron counting rate. When the radium source was brought to within 10 cm, a change of approximately 10 per cent in the counting rate of the neutron counter was noted.

---

[15] S. A. Korff, *Rev. Sci. Inst.*, 12, 94 (1941).

It will be recalled that the number of neutrons will not be materially diminished by 8.5 cm of lead, whereas the gamma rays will be substantially reduced. It was estimated that the gamma ray flux through the counter due to the radium-beryllium source inside the lead block was approximately 0.2 times that encountered in the stratosphere. With the additional radium source at 1 meter, the gamma ray flux was approximately double that in the stratosphere, while with the radium at 10 cm, the gamma ray intensity passing through the counter was some 200 times the stratosphere intensity. It was noted that no appreciable change in the counting rate was produced by the gamma rays until the gamma ray flux through the counter was many times that which the instrument would encounter in the stratosphere. It appears reasonable, therefore, to conclude that the gamma ray intensity in the stratosphere does not account for an appreciable number of the discharges of the apparatus nor does this intensity produce a decrease in the efficiency [16] of the detecting unit. In this case the counter was operated at a potential several hundred volts below the Geiger threshold. At higher voltages the effect due to gamma rays is more marked.

In the neutron detecting unit used in connection with the neutron-shower coincidence measurements, it would require a shower of approximately 1000 electrons simultaneously passing through the counter to create as much ionization as is produced by an alpha particle resulting from a neutron-induced boron disintegration. The counting rate of this counter due to large showers will, therefore, be that due to the total number of 1000-ray showers passing through the counter. Inasmuch as the lead and much of the paraffin was located below the counter, and since showers in the main go downward, it does not seem probable that spurious counts due to 1000-ray showers can account for the total number of observed counts. It should be added, however, that the shower tray under the lead would almost certainly discharge in the event that such a shower took place in the matter above it.

Further controls were used, to make certain that none of the observed coincidences were due to pulses coming in to the apparatus from the voltage supply line, or to spurious effects induced by one part of the set on the other. This was ascertained by operating the set under normal conditions except that the voltage on the BF_3 counter was reduced just below that necessary to produce counts. The amplifiers remained sensitive and might be expected to record pickup or induced effects if any existed. No counts were observed in a 20 hour run.

---

[16] C. G. Montgomery and D. D. Montgomery, *Rev. Sci. Inst.*, **11**, 237 (1940).

Finally, in order to establish that the coincidence counts themselves were due to slow neutrons, a cadmium shield was placed over the neutron counter when it was running, and known to be counting coincidences properly. No coincidences were recorded in 22½ hours with this arrangement. This sets an upper limit to the counts produced by accidental, large showers, and fast neutrons.

D. Possible Sources of Neutrons

Inasmuch as neutrons are not thought to be primary particles for reasons of their instability and other considerations referred to above, we must consider the possible processes by which they may be produced by the cosmic radiation in the atmosphere. Any such process must fulfill three requirements. (1) It must give the correct altitude dependence. (2) It must account for the observed coincidence between neutrons and showers. (3) It must be a process with a reasonable cross section. On the basis of these three requirements, we may analyze several possible sources of the neutrons.

1. *Neutrons Produced by Primary Protons*

If we assume that 10 per cent of the total primary flux of cosmic rays are protons, then approximately 1 proton per square cm per minute arrives at the upper limit of the atmosphere. This current I of protons must produce the observed $N = 100$ neutrons per cm sq. per minute in the first meter of water equivalent below the top of the atmosphere. The production cross section (σ) will be given by:

$$N = I\sigma M,$$

where M is the number of producing centers (nuclei) in the first meter of water. If we take $M = 2.7 \times 10^{24}$, then $\sigma = 4 \times 10^{-23}$ square cm, a cross section which appears to be unreasonably large.

2. *Neutrons Produced by any Primaries*

If we suppose that the neutrons are produced by any primary cosmic ray regardless of its nature, the cross section calculated above will evidently be reduced by a factor of 10. This leaves it still too large to represent a process which one would, on the basis of present knowledge, reasonably anticipate.

3. *Neutrons Produced by Photons*

If we suppose that the neutrons are produced by photons of more than say about 10^7 volts which are present in the radiation, then at a depth of one meter of water, I will be 1000 and the cross section will be 4×10^{-26} square cm. This cross section, while somewhat larger than that for gamma-ray photoneutrons, does not appear to be unreasonably large, especially if one considers that the cross section may increase with increasing photon energy above 10^7 volts.

Further information may be gained by considering the altitude dependence of the neutron intensity. We have shown that the intensity increases with altitude at the same rate as does the soft component. It will be recalled that neutrons are essentially local phenomena and seldom diffuse more than 1 meter of water equivalent from the place where they are produced. Hence we would expect the intensity of neutrons to depend upon altitude in the same manner as does the producing agent. The neutron intensity as a function of elevation is shown in Fig. 3. At the low elevation the altitude dependence of the total intensity is quite different from that of the neutrons, suggesting that most of the neutrons are probably not produced by mesotrons which constitute most of the cosmic ray intensity at sealevel. At the higher elevations where the total intensity is chiefly produced by photons and electrons, the altitude dependences of the neutrons and of the total intensity are parallel.

E. Coincidences

Finally the observed rate of coincident discharges obtained with the neutron counter and the shower tray indicates that there is a connection between these processes. If the neutrons were produced chiefly by photons, then we should expect just such a connection. There will, of course, be more photons present in the larger showers and the coincidences indicate that there are more neutrons associated with the larger showers. Since it has been shown that the neutrons do not carry a large fraction of the total cosmic ray energy,[7] this process will not interfere with the normal cascade production of showers.

We may, therefore, conclude that while any cosmic ray particle can, energetically speaking, produce neutrons, it appears probable that most of the neutrons are produced by the photons and electrons in the soft component. Cosmic rays thus provide a new line of evidence for a high-energy process, the photodisintegration of nuclei.

ACKNOWLEDGMENTS

The author wishes to express his thanks to the American Philosophical Society for a grant which enabled him to obtain a neutron source used in the above experiments. He is also indebted to the Carnegie Institute of Washington for financial support in connection with the balloon flight program. Thanks are due to the Mount Washington Observatory for providing facilities and coöperating with the installation on Mount Washington, N. H., of the neutron measuring apparatus. The author also wishes to express his appreciation to E. T. Clarke who assisted with many of the observations, and to his colleagues at the Bartol for helpful discussions.

STRUCTURE AND DEVELOPMENT OF CENTRIFUGED EGGS AND EARLY EMBRYOS OF DROSOPHILA MELANOGASTER

RUTH B. HOWLAND [1]

Department of Biology, Washington Square College of Arts and Science,
New York University

ABSTRACT

The present paper offers data on the cytology of *D. melanogaster* eggs and early embryos after the application of a centrifugal force strong enough to stratify the egg constituents but not so drastic as to inhibit further development. Certain consistently occurring structural variations found in larvæ, pupæ and in the adults hatching from centrifuged eggs are discussed in reference to the early histological pictures. The results are aligned with those obtained by Hegner (1909) in Chrysomelid beetles, by Pauli (1927) on Calliphora and Musca, and by Reith (1931 and 1932) on Campanotus and Lasius.

I. MATERIAL AND METHODS

DROSOPHILA melanogaster eggs were obtained from the Oregon R stock long maintained at Washington Square College of Arts and Science. The stocks were kept in the laboratory incubators at 25° C.

Fifteen pairs of four-day males and females were transferred from bottles, in which active laying had occurred, to vials containing yeasted banana-agar slabs. At the end of half an hour, the first egg-laying slabs were discarded and replaced by new. From these, eggs were collected at known intervals for centrifugation.

Eggs in groups of twenty-five were carefully oriented, one at a time, in small, deep holes dug in soft, moistened blotter or in stiff agar disks. These were transferred to flat bottomed vials, which were wrapped in cotton and wedged into centrifuge tubes. Speeds and time of centrifugation were varied over a wide range (from 5 or 7 minutes at 3000 r.p.m. to 30 minutes at 2400 r.p.m.). The blotters or agar blocks were kept at a uniform distance from the center of revolution, and the same centrifuge (Size No. 1, Type SB, manufactured by the Independent Equipment Company) was used throughout the work. Some swerving from the desired orientation was noted, and this was checked when the eggs were dechorionated after centrifugation. If subsequent development

---

[1] Aided by a grant from the Penrose Fund of the American Philosophical Society of Philadelphia.

were to be followed through to eclosion, centrifuged dechorionated eggs were classified into groups according to the angle of stratification of the egg constituents. Orientation with the anterior end centrifugally placed is referred to as anterior centrifugation; with posterior end outward as posterior centrifugation. The same descriptive terminology is used for dorsal and ventral orientation. Oblique stratifications are spoken of as antero-lateral, postero-lateral, etc.

For studying the behavior of living centrifuged eggs, Belar's solution was found to be a satisfactory immersion medium. Eggs to be sectioned were pricked in FAA solution, and after the usual procedures of embedding, sectioned ar 8μ and stained in Heidenhain's Iron Hæmatoxylin, Delafield's Hæmatoxylin, or Harris stain.

II. Constituents of the Early Egg

As in many telolecithal eggs in which the position of the egg nucleus is eccentric, the centrolecithal egg of *Drosophila melanogaster* shows a definite polarity and a certain axial distribution of the inner egg-materials in the normal undisturbed condition.

At the anterior pole, in granular ooplasmic islands which are practically yolk-free, the egg and sperm nuclei lie. Backward along the long axis the yolk, consisting of spherical bodies of varying sizes within vacuolar areas, becomes somewhat more concentrated toward the posterior pole. The clear, non-granular, outer cortical ooplasm is more or less uniformly distributed over the egg surface. It occurs in slightly greater concentration just under the vitelline membrane at the anterior pole, and at the posterior pole where it has the form of a thickened cap (pole-plasm) heavy with definite but fine granules, once thought to be germ cell determinants.

III. Experimental Findings

1. Stratification of Eggs in Early Cleavage Stages.

Early eggs centrifuged for thirty minutes at 2400 r.p.m. (Pl. I, Fig. 1) are stratified into five zones.[2]

(*a*) *White Zone.* Centrifugally, a layer of heavy white material, which is non-staining with Iron Hæmatoxylin, Delafield's

---

[2] The description refers only to fixed and stained eggs.

Hæmatoxylin and eosin, or Harris' stain. This is formed by the confluence of the white vacuolar material which in the normal egg surrounds the heavier yolk bodies or, more rarely, occurs as clear spherules. The marginal borders of this zone may be wavy or crenulate as if the vacuolar fusion had not been quite complete when reached by the killing agent.

(*b*) *Yolk Zone.* Next in weight are the yolk bodies released from their surrounding white fluid areas. These mass together so closely as to crowd out other cytoplasmic elements. A gradation is seen within this zone; the larger, heavier granules bordering the white zone just described.

(*c*) *Granular Zone.* A very broad zone consisting of dark granules of uniform size and staining quality. These on closer examination, especially in regions of puncture (for fixation), give the appearance of a fine fibrillar network. The fibrillar fixation patterns show parallel or swirled configurations, or "stress lines" near the puncture point. At times, free torn fibrillae mark the region where the needle was withdrawn. The egg nuclei lie in this zone, and fibrillae follow the contour of their walls.

(*d*) *Pale Zone.* Centripetally, a pale zone is distinctly separate from the third or granular stratum. It is more homogeneous in appearance and under oil immersion uniformly distributed extremely fine granules can be seen. In torn sections of fixed eggs, it lifts off intact from the underlying fibrillae, the edges of which "fray out" along the line of separation. In these early stages there are no nuclei in this zone, though spindles and resting nuclei frequently lie against its centrifugal border. No fibrillar patterns are visible. It is suggested that this zone may consist partly of cortical ooplasm.

(*e*) *The Pole-plasm.* In addition to these zones, there is one already differentiated layer in the newly laid egg which is moved slightly, if at all, from its normal position, and which if it does shift, moves as a whole. This is the posterior pole-plasm with its granules, which because of its continued integrity as a much more strongly basophilic area, and its fixed position at or near the posterior pole, may be designated as a fifth zone.

At the lower centrifugal speeds (10 minutes at 1600 r.p.m.) the pole zone, *d*, may not be distinct from the granular zone, *c*. The other zones are as described above, the elements of each being less densely packed together.

2. *Nuclei in Early Stages.*

Early cleavage stages, centrifuged for 30 minutes at 2400 r.p.m., may show a giant polar-body complex with multipolar spindle fibers on which are massed large numbers of chromosomes (Pl. I, Fig. 2), or the three polar bodies may be separated from each other and pressed with such force against the vitelline membrane that they come to lie in small outpocketings, and thus strongly resemble the typical protruding polar bodies of eggs of other animals.

Cleavage nuclei are found in different stages of mitosis. The usual synchronous division rhythm has been disturbed by the continuous force applied (Fig. 2). Giant cleavage nuclei are common in which there may be twice to three times the normal amount of chromatic material. Lagging chromosomes are found on the telophase spindles. Not all spindles lie parallel to the axis of the force applied. Occasionally spindles may even lie at right angles to this force (Pl. I, Fig. 2). All nuclei remain in the "fibrillar" zone, *c*.

3. *Stratification in Preblastoderm Stages.*

In preblastoderm stages (ranging from 2 to 3 hours at 25° C.), centrifuged from 6 minutes at 1200 r.p.m. to 7 minutes at 3000 r.p.m., the zonation differs from that described in earlier eggs in one important feature. Although in the egg center it is often possible to recognize four zones, *a*, *b*, *c*, and *d*, a layer of superficial cortical ooplasm (keimhaut) and the nuclei which have come to lie there preliminary to blastoderm formation, do not shift from their position. The greater part of the egg surface is covered with the keimhaut in which the preblastodermic nuclei have been held even in the absence of cell boundaries. Such a condition must be due to an increase in viscosity or an added "rigidity" of the cortical material on arrival of the cleavage nuclei.

Nuclei (vitellophags) which are not "bound" by the cortical ooplasm aggregate at the usual nuclear level, *i.e.*, within the "fibrillar" or heavy granular zone *c*, and may be so numerous here as to form a compact tissue-like mass.

In 45–60 minute eggs, after posterior centrifugation and a subsequent recovery period of one hour, the cortical ooplasm is of unequal thickness at different levels along the longitudinal axis. Nuclei in the thinner layers are small, in thicker layers large, and although not separated by cell walls they are in different stages of mitotic division. Preblastodermic nuclei are frequently absent at the anterior end. The internal nuclear mass is especially large and may form a solid core at the "fibrillar" level. In some the large yolk platelets are packed against a posterior blastoderm, in others against the vitelline membrane from which zone *e* has shifted slightly to the right or left of the midline as a compact unit. (Note discussion of possible causes of rotated abdomen in reference to this type of stratification in later section of paper.)

4. Larvæ from Centrifuged Eggs.

(*a*) The hatching time of larvæ from centrifuged eggs is invariably longer than normal, but varies with many factors such as age and orientation of the egg, and degree of dislodgment of materials at the time of centrifugation. As would be expected, the younger the egg, the greater the delay. In general, a larger percentage of early eggs hatch after either dorsal or posterior than after ventral or anterior centrifugation. The larvæ which fail to pupate but continue for many days to crawl and feed have doubtless suffered some injury to Weismann's ring, which has prevented the secretion of the pupation hormone (Hadorn and Neel, 1938). Of the hatched larvæ which do not go through to pupation, many appear normal but sluggish, and others have tracheal tubes in a snarled mass postero-dorsally. A marked torsion or partial rotation of the body along the antero-posterior axis which will also be described in the adult fly, may be observed in the larval stages (Pl. I, Fig. 3).

(*b*) Unhatched larvæ, 48–63 hours of age at 25° C., have also been observed in active movement within the egg membranes. Such delayed larvæ developed from eggs centrifuged anteriorly, antero-ventrally, and antero-dorsally, very frequently have malformations in the jaw armature (Pl. I, Fig. 3). Common variations in jaw structure are knobbed terminal prongs, scattered chitin bars, and clumped chitinous masses. These malformations obviously preclude the normal hatching procedure leading to the rupture of the opercular membrane. Sections show a well developed proventriculus in almost every animal. Mid and hindguts in old unhatched larvæ are often greatly distended as if digestive fluids had been secreted in quantity. The almost total absence of yolk in these larvæ together with the distended distal ends of the gut cells, a common histological sign of secretory activity, strengthen this interpretation.

5. Pupæ.

Torsion is strongly marked in puparia formed by larvæ in which rotation of the body on the long axis has resulted from centrifugation (Pl. I, Fig. 4). The posterior tracheal tubes may be laterally rotated (either clockwise or counter-clockwise) in respect to a median bilateral plane passing through the head and thorax. Deletions of one or more tracheal endings are also common.

6. Adults.

The greater percentage of flies (an average of 60–70 per cent) hatching from centrifuged eggs are normal in all external characters. This may also be true after centrifugation at much higher forces (Shapiro, 1938). As a rule, the older the egg when centrifuged, the higher the percentage of normal flies (except in cases of extremely prolonged or violent centrifugation which produces ruptures).

The various adult abnormalities observed fall into two main categories:

(*a*) *Deletions or Dislocations.* Right or left antennæ are frequently missing. In three cases, 3-hour eggs centrifuged for 20 minutes at 1600 r.p.m. showed, on dechorionation, the heavy white zone and heavy yolk thrown antero-ventrally. From these three eggs, adults hatched lacking one antenna. External genitalia may be absent entirely, or the genital plates may be present inside of the abdomen (as in the mutant fly, *abdomen rotatum*). They may be recovered by dissection and identified.

Dorsal chitinous abdominal plates often fail to meet in the midline, only a thin chitinous membrane bridging the gap. Ventral bristle tufts may not be joined in one median zone, but remain separated as right and left bristle areas.

(*b*) *Torsions.* From rotated puparia already described as formed by twisted larvæ, hatch adults with varying degrees of body torsion (Pl. I, Fig. 5). Observed twistings were largely abdominal, and were as often clockwise as counter-clockwise. Torsions vary through all degrees of rotation up to a 180° displacement. That they are not caused by lateral tensions exerted through deletions of dorsal abdominal plates is evident, for the dorsal chitin may be normal both in extent and placement. On following through individual protocols on the development of dechorionated eggs, a high correlation is observed between obliquely centrifuged eggs (postero-lateral centrifugation) and rotated adults. Though originally oriented with posterior poles outward and long axes parallel to the force applied, eggs often slip or shift to right or left during centrifugation. When dechorionated, the heavy white material (of zone *a*) and yolk (of zone *b*) are seen to be packed terminally against the vitelline membrane either to the right or left of the midline. These masses of heavy material lie in such a position that, if not later dislodged or redistributed, they doubtless serve as mechanical obstructions in the subsequent elongation of the germ band (see discussion).

IV. Discussion

The distribution of the chief mass of materials in centrifuged Drosophila eggs conforms in the main to that described by Pauli (1927) in eggs of Calliphora and Musca, and Hegner (1909) in the eggs of the Chrysomelid beetles. Both observed after moderate centrifugation (1) a centrifugal yolk zone, (2) a central plasma zone with cleavage nuclei, and (3) a light zone of fatty material. Pauli also found after prolonged centrifugation, that the yolk showed its dual composition by separating into two strata, one on each side of the plasma zone.

In Drosophila the heavy yolk bodies become separated from their still heavier surrounding fluid and the two stratify as sharply contrasted layers. It is probable, therefore, that we are also dealing here with two kinds of yolk. This point will be considered again later in the discussion.

The behavior of the cortical material (keimhaut) of insect eggs on centrifugation has been variously described. Hegner stated that on weak centrifugation some of this material was thrown centripetally, while still stronger forces piled up the entire keimhaut at the centripetal pole. Pauli also found that the keimhaut after moderate centrifugation was thicker at the centripetal end. After prolonged centrifugation, it was entirely lacking from the centrifugal end, nor could it now be identified at the centripetal pole occupied by the fatty material. She reported also that if cleavage nuclei had not formed a blastoderm (by migration into the cortical ooplasm) before being subjected to a long period of centrifugation, they would lay down a blastoderm while the force was being applied, but only at the centripetal surface of the "plasma zone." If, on the other hand, the nuclei had already reached the keimhaut before centrifugation, the entire plasma region was found to have a well formed blastoderm. Though she did not call attention to the holding quality of the cortical area subsequent to nuclear arrival, evidence is here indirectly given to substantiate a profound change in its viscosity.

Reith (1931, 1932) in his experiments on the ants Campanotus and Lasius, found that the degree to which development will proceed after centrifugation is linked with the behavior of the cortical ooplasm. If young eggs were centrifuged prior to the zonation and localized thickening of the cortical ooplasm—a process here controlled by the distribution of some non-stainable material from a posterior activation center—the stratification of egg materials prevented further development. On the other hand, once the cortical ooplasm had streamed into its normal position and pattern, development proceeded after centrifugation, and germ bands formed. In the Drosophila egg, we have no evidence of interaction between an activation center and the remaining ooplasmic material. At the moment of fertilization, a clear cortical zone is present, slightly thicker at each egg pole. Evidence of some displacement of this material by centrifugation is seen in the early cleavage stages. However, if the cleavage nuclei have already reached the cortical layer before centrifugation, though cell walls have not as yet separated them and established individual blastoderm cells, the increased rigidity and holding quality of the surface layer is very striking. Nuclei, once arrived, are not shifted from their surface position. As would be expected, later eggs show a much greater hatching percentage. All these data emphasize the essential role played by the cortical material in the establishment of the first development pattern and suggest that this cytoplasmic layer and the cleavage nuclei so react as to form a viscous, syncytial surface layer, or preblastoderm, in which, subsequently, localized changes occur.

One region of the Drosophila egg, the pole-plasm[3] with its granular contents (zone *e*), varies in behavior from the other ooplasmic constituents. Even in the earliest stages centrifuged, this terminal cap retains its identity as a more rigid area, and though shifted slightly from its original position, is not broken down, nor are its granules displaced. Similar results were obtained by Hegner (1909) who reported that the entire mass of polar granules in the beetle egg moved into the central plasma layer as a unit when the posterior end of the egg was placed centripetally. The granules did not disperse after removal from the centrifuge, although other egg constituents were redistributed. In Calliphora and Musca, Pauli was unable to establish the behavior of the pole plasm and its granular inclusions.

---

[3] Really a part of the cortical ooplasm.

Though we have no evidence of polar material streaming through the Drosophila egg, an interesting parallel still obtains between this obviously specialized area and the activation center ("Bildungszentrum") of the odonate egg (Seidel, 1932). Entrance of cleavage nuclei into this center in the Platycnemis egg initiates the streaming forward of materials quite essential to further chemo-differentiation while in the dipteran egg the entrance of cleavage nuclei (or a single nucleus, Kahle, 1908) into a similarly placed region is an essential precursor to the softening and bulging of the terminal membrane and setting aside of primordial germ cells. Both, then, have posterior areas in which chemical changes are brought about by the arrival of cleavage nuclei.

Some points of interest arise in considering the so-called "white" zone (*a*) of Drosophila described in a previous section of the paper. The zone is derived from the colorless and non-staining material originally surrounding the yolk spherules or from occasional clear vacuolar areas. This colorless part of the yolk "vacuole" has been considered by Huettner (1923), from a study of stained slides, to be a shrinkage area or histological artefact. Rabinowitz (1938) also favors this interpretation, but adds that these areas "may correspond to the large vitelline spheres described by Nelson in the egg of the honey-bee." Hasper (1911) describing the yolk content of the Chironomous egg speaks of easily soluble, colorless fat drops, the refractive index of which in contrast to that of the yolk spheres is only slightly higher than that of water, and Noack (1901) distinguishes in the Calliphora egg two yolk zones—a primary plasma-rich peripheral yolk zone and a plasma-poor yolk zone with large vacuoles partly confluent. Present findings strengthen the evidence that the material in the Drosophila egg is similar to the colorless fatty material of Chironomous, or the vacuolar contents of Calliphora. Special killing and dehydration methods followed by differential staining may later determine whether it is of a fatty nature.

With the picture of stratification of egg constituents in mind, there remains to be considered the relationship between these imposed nuclear and cytoplasmic disturbances and the normal or abnormal conditions later found in Drosophila larvæ, pupæ, and adults. Evidence of isopotentiality of early cleavage nuclei is clear in that displacement by force, through altering the usual course of nuclear migration, does not prevent the formation of normal larval organs and hatching of normal adults. Rupture, during centrifugation, and loss of some material has been observed in preblastodermic stages. This would easily account for the occasional deletion of an organ (antenna of an adult). In fact, such a condition would be in complete alignment with earlier experimental investigations involving the removal, by pricking of egg materials (Howland and Child, 1935) and the subsequent absence or malformation of adult organs or regions.

To interpret the larval, pupal and adult torsions, however, is a more difficult matter. At first glance it would seem that, during centrifugation, the entire terminal egg mass had swung either clockwise or counter-clockwise in reference to an antero-posterior longitudinal axis, and remaining in this position, gone on to form the larval and adult organs usually laid down there. Such an explanation rests on the false assumption that gravitational pull acting as a counterforce during centrifugation is of sufficient strength to shift the heavier egg constituents downward as they stratify centrifugally.

The study of different stratification patterns in eggs dechorionated after centrifugation and grouped in various categories on the basis of yolk position offers another and sounder suggestion. It was observed that a large percentage of obliquely stratified (postero-lateral yolk mass) eggs gave twisted larvæ and pupæ which eclosed as "rotated" flies. Normally, the lengthening germ band follows the curved posterior end of the egg in the median line and travels along this line mid-dorsally. On contraction at 10 to 12 hours it follows the same course in reverse, bringing the anus to its postero-terminal position. But if, on extension, the germ band were deflected obliquely by a heavy firm mass of centrifuged yolk, contraction would bring the anus into a lateral position and coincident with this the terminal abdominal segments would be rotated.

Attendant conditions often found—as for example internal genitalia—might also result from adhesions and infoldings of cell groups along the surface of tightly packed yolk masses.

The striking similarity between rotated flies resulting from centrifuged eggs, and the Drosophila mutants characterized by rotated abdomens (ar/ey^D of Baliajeff 1931, and others) is obvious. In the mutant stocks, however, the puparial case shows no asymmetry, nor do the larvæ appear to be asymmetrical (Marengo 1941, unpublished).

V. SUMMARY

1. Early eggs of Drosophila melanogaster centrifuged for five minutes at 3000 r.p.m. stratify into five zones: (*a*) heavy white zone (probably yolk material) centrifugally, (*b*) zone of yolk spherules graded among themselves from large to small, (*c*) granular protoplasmic zone with nuclei, (*d*) pale centripetal zone partially made up of cortical ooplasm, and (*e*) the posterior pole plasm really a portion of the posterior cortical ooplasm with its contained granules.

2. The pole plasm, (*a*) shifts only slightly, if at all, from its original terminal position. It maintains its integrity as a strongly basophilic heavily viscous area from which the granules are not displaced.

3. Centrifugation of early eggs may cause a breakdown in the normal synchrony of nuclear division. Giant nuclei are often found.

4. The viscosity of the cortical ooplasm shows a marked change on arrival of cleavage nuclei within it. Preblastodermic nuclei are not dislodged by centrifugation, even though cell walls are not present when the force is applied.

5. The hatching time of larvæ is increased and many fail to hatch because of jaw abnormalities. Torsion is common in hatched larvæ.

6. From rotated larvæ rotated puparial cases are formed.

7. Deletions, dislocations, and abdominal rotations are common abnormalities in flies hatched from centrifuged eggs. A large percentage of the flies which hatch are normal in all external characters.

VI. LITERATURE CITED

BALIAJEFF, N. K. 1931. Erbliche Asymmetrie bei Drosophila. Ein neues Gen im IV chromosom von D. melanogaster. *Biol. Zentr.*, **51**: 701–709.

HADORN, ERNST AND JAMES NEEL. 1938. Der hormonale Einfluss der Ringdrüse (Corpus allatum) auf die Pupariumbildung bei Fliegen. *Arch. f. Entw-Mech.*, **138**: 281–304, Heft 2.

HASPER, MARTIN. 1911. Zur Entwicklung der Geschlechtsorgane von Chironomus. *Zool. Jahrb.*, **31**: 543–602.

HEGNER, R. W. 1909. The Effects of Centrifugal Force upon the Eggs of Some Chrysomelid Beetles. *Jour. Exper. Zool.*, **6**: 507–552.

HOWLAND, R. B. AND GEORGE CHILD. 1935. Experimental Studies on Development in *Drosophila melanogaster*. I. Removal of Protoplasmic Materials during Late Cleavage and Early Embryonic Stages. *Jour. Exper. Zool.*, **70**: 415–427.

HUETTNER, A. F. 1923. The Origin of the Germ Cells in *Drosophila melanogaster*. *Jour., Morph.* **37**: 385–423.

KAHLE, W. 1908. Die Pædogenese der Cecidomyiden. *Zoologica*, H. 55, 1–80.

MARENGO, N. 1941. Some Developmental Aspects of the Drosophila Mutant "Abdomen rotatum." (Personal communication.)

NOACK, W. 1901. Beiträge zur Entwicklungsgeschichte der Musciden. *Zeit. f. wiss. Zool.*, **70**: 1–57.

PAULI, M. E. 1927. Die Entwicklung geschnürter und centrifugierter Eier von Calliphora vomitoria und Musca domestica. *Z. wiss. Zool.*, **129**: 483–540.

RABINOWITZ, M. 1938. Studies on the Cytology and Early Embryology of the Egg of the *Drosophila melanogaster*. Unpublished. Ph.D. thesis at Washington Square College.

REITH, F. 1931. Versuche über die Determination der Keimesanlage bei Camponotus ligniperda. *Z. wiss. Zool.*, **139**: 664–734.

——. 1932. Ueber die Localisation der Entwicklungsfactore im Insektenkeim. 1. Zentrifugierversuche an Ameiseneiern. *Arch. f. Entw.-Mech.*, **127**: 283–299.

SEIDEL, F. 1932. Die Potenzen der Furchungskerne im Libellenei und Ihre Rolle bei der Aktivierung des Bildungszentrum. *Arch. f. Entw.-Mech.*, **126**: 213–276.

SHAPIRO, H. 1938. The Embryogenesis of Drosophila following Centrifuging at High Forces. *Anat. Rec.* **72**: 101–102 (Suppl.).

EXPLANATION OF PLATE I

FIG. 1. Young egg, centrifuged antero-ventrally, dechorionated and sectioned longitudinally. Stratification oblique. Five zones easily identified are: (*a*) white centrifugal zone, (*b*) yolk zone, (*c*) granular zone in which the fibrillar pattern is especially visible near the puncture point (*p*), (*d*) pale zone, and (*e*) zone of pole plasm with its characteristic granules. Note that zone (*e*) has not moved from its normal position. The micropylar region (*m*) of the vitelline membrane is prominent.

FIG. 2. Composite diagram of the anterior end of an 8-celled cleavage stage in which stratification was oblique. Two zones are shown (*d*) the centripetal pale zone, and (*c*) the granular zone in which the polar body complex (*p.b.c.*) and eight cleavage nuclei (*n*–1 to *n*–8) are found lying at various angles and in various phases of mitosis. Nuclei 1, 2, 3, 4, 6, and 7 are in prophase; nucleus 8 is in anaphase, and nucleus 5 is in telophase (heterochrony). No nuclei have lodged in the pale centripetal zone.

FIG. 3. Unhatched rotated larva (or late embryo) from centrifuged egg.[4] Jaw armature (*j*) formed, but with terminal hook detached (*h*).

FIG. 4. Puparium showing rotation of terminal region. Posterior tracheal openings (*tr*) are lateral.

FIG. 5. Adult with rotated abdomen (*r.a.*) from centrifuged egg.

[4] Acknowledgment is made to Dr. Harry Albaum of Brooklyn College for this sketch. Dr. Albaum and Dr. Samuel Kaiser of Columbia University are interested in another phase of this work.

PLATE I

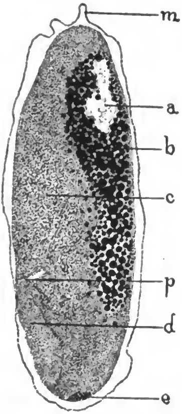


FIG. 1

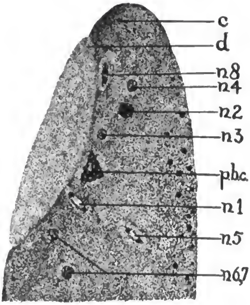


FIG. 2

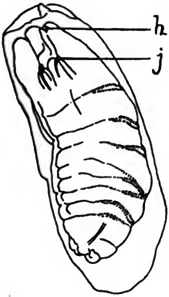


FIG. 3

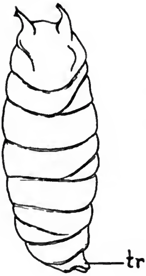


FIG. 4

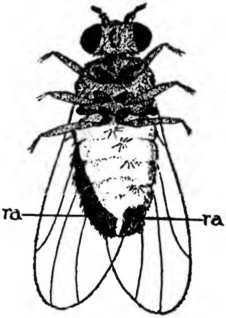


FIG. 5

EXPERIMENTAL STUDIES ON THE REPRODUCTIVE PHYSIOLOGY OF THE MALE SPRING PEEPER, HYLA CRUCIFER

ROBERTS RUGH*

Washington Square College of Arts and Sciences, New York University

Abstract

This study is a description of the morphology and physiology of the genital system of the Spring Peeper, *Hyla crucifer*, and includes that portion of the excretory system with genital functions. Taking advantage of the fact that the anterior pituitary hormone will cause release of spermatozoa from the testes of the hibernating frog, this method was used to trace the path of sperm from the seminiferous tubule, through the vas efferens and into the kidney. Within this organ it was found that of the 400 or more malpighian corpuscles, about 10 per cent (all located in the anterior third of the kidney) are seen to contain spermatozoa during active breeding, and the malpighian corpuscle represents the point at which the genital and the urinary ducts converge. Spermatozoa are seen abundantly within the 35–40 malpighian corpuscles, and passing outwardly in the uriniferous ducts (urino-genital ducts) through the kidney substance to the ureter. There is no morphological or histological evidence of a Bidder's canal (supposedly passing along the mesial margin of the anuran kidney) which might carry spermatozoa to the ureter at the posterior tip of the kidney (as described for European forms) and it is shown that the presence of such a canal is functionally unnecessary. The role of the seminal vesicle is more important than the same structure in other frogs, since in *Hyla crucifer* (as in toads) each egg emerges from the female cloaca singly and must be inseminated singly. In general, however, the genital morphology and physiology of *Hyla crucifer* is quite frog-like, the major differences relating to the large malpighian corpuscles in the reduced kidney, as compared with *Rana pipiens*.

The general morphology of the various anuran urino-genital systems has been described by Spengel (1876), Kolliker (1875), Neumann (1875), Nussbaum (1886) and Ecker (1889) on European frogs and more recently by Hall (1904) on *Rana sylvatica*, Filatow (1904) on *Rana esculenta* and Gray (1930) on *Rana temporaria*. Studies on the vasa efferentia were made on some Australian species by Georgiana Sweet (1908); on the sexual portions of the kidney by Bidder (1846), Wiedersheim (1886), Marshall & Bles (1890), Hall (1904) and Gray (1930). The structural relationship of the Bidder's canal to the genital ducts was described by Beissner (1898) and its evolutionary significance was suggested by Ponse (1927). It was Hyrtl (1863) who first suggested the sex function of the malpighian corpuscle and Noble (1931) stated that the smaller corpuscles alone had the dual function of urine and sperm transport.

---

* Aided by a Grant from the Penrose Fund of the American Philosophical Society of Philadelphia.

As suggested by Gray (1930) and Noble (1931) it is a very dangerous practice to assume similar morphology and physiology even for closely related forms. The presence, for instance, of a Bidder's canal running along the mesial margin of the anuran kidney and joining the mesonephric duct (ureter or Wolffian duct) at the posterior extremity of the kidney, has been assumed for all American species. Such a canal was not found in *Rana pipiens* (Rugh 1939).

Through experimentally induced sexual activity it is possible to trace the sex cells from their gonads to their destination. If female amphibia are studied at appropriate intervals after pituitary treatment the course of eggs (Rugh 1935) and of spermatozoa (Rugh 1939) can be followed through the entire urino-genital system. The findings by this method have led to certain modifications of earlier morphological considerations, later corroborated by micro-injection of india ink into the tracts themselves.

It is the purpose of this study to point out the major differences in reproductive physiology of the male *Hyla crucifer* as compared with the larger and more common leopard frog, *Rana pipiens*.

Materials and Method

The subject of this investigation was *Hyla crucifer*, variously known as Pickering's Tree Toad, Pickering's Hyla, Peeping Frog, Castanet Tree Frog, and the Spring Peeper. While some of its breeding reactions may be more toad-like, general morphological considerations place these animals among the frogs. Specimens were caught in their native habitats during the latter part of March, the time when they normally emerge from hibernation. *Hyla crucifer* lives in lowland swamps, generally wooded and grassy. Its small size makes it a very difficult animal to find, the males measuring 18–30 mm and the females 20–35 mm in body length. During early spring, particularly at night, numerous males may be heard issuing their shrill, high-pitched and clear call. At such times the male throat pouch is seen to enlarge to a size greater than the head of the animal itself. The males squat on twigs and leaves at the edge of the pond, usually under some overhanging object, ready to leap into the water.

The male Hyla is usually dark brown in color, although there is great color variation even among specimens caught at the same time and in the same locality. The females are always lighter in color than the males and show the characteristically bulky abdomen of female frogs. In the lighter colored specimens of either sex there is visible on the dorsal surface dark bands which form a large cross, and additional bands may be seen on the appendages. The snout is definitely pointed, projecting somewhat beyond the lower jaw. The tips of the fingers are provided with adhesive pads by means of which these animals can cling even to the sides of a smooth glass aquarium. The fingers are not webbed. Folds in the skin and a definite darkening beneath the jaw of the male distinguishes it from the female.

The normal breeding period of *Hyla crucifer* is from April to June, although in the Southern States this same species is known to breed as early as February. Temperature is undoubtedly the major factor in determining the breeding season. The eggs are very small, and of a light brown color. A single female may produce as many as 1,000. The eggs are generally laid at night, with pairs of animals floating at or near the surface of the water. Eggs are liberated from the female singly and as the egg is dropped the male clamps its cloaca over the point of emergence of the egg, liberating spermatozoa.

The specimens used in this study were caught a few weeks before they would have become sexually active, *i.e.,* just as they emerged from hibernation. They were segregated into pairs of the same size, caught at the same time and in the same environment. One was used as a control and the other member of each pair of males was used as the experimental animal, to be treated with the anterior pituitary hormone. The glands injected were excised from adult specimens of the leopard frog, *Rana pipiens,* and two female glands were injected into the body cavity of each of the experimental animals. From 9–36 hours after injection, various animals were killed and their entire urino-genital systems were fixed in Bouin's fluid. The controls were injected with a similar amount of frog brain tissue unrelated to the pituitary; were kept in the same aquarium as the experimental animals; were killed simultaneously with them; and their organs were similarly fixed. All tissues were embedded in a mixture of 90 per cent—58° C. paraffin; 5 per cent bayberry wax and 5 per cent beeswax. This mixture had a melting point of about 55° C. and allowed sectioning at almost any thickness and at any time of the year without resorting to the use of ice and freezing of the microtome knife. Sections were cut at 10 μ and the entire urino-genital system of a single specimen was mounted serially. The stain used was Heidenhain's Iron-alum and Haematoxylin with no counter stain. Destaining for nuclei brought out the spermatozoa beautifully.

Observational and Experimental Data

The Testis

The testis of *Hyla crucifer* is completely blackened by pigment which not only surrounds the organ, beneath the peritoneum, but which also encapsulates each of the seminiferous tubules (Pl. I and Pl. II, Fig. 1). The average testis measures about 2.15 mm in length and has a thickness of about 1.10 mm. It is made up of about 150–200 oval shaped seminiferous tubules, each morphologically distinct from its neighbor. Interstitial tissue is negligible and trabeculae cannot be identified, but each tubule is encapsulated by a layer of pigment. This pigment is lacking in all other frog and toad testes studied. It would be difficult to determine whether there is any change in this small amount of interstitial tissue associated with breeding reactions. Within the circle of connective tissue which surrounds each tubule there is a single layer of flattened epithelial cells. Next may be seen conically projecting masses of cells which are probably the spermatogonia held over for the following season (Pl. II, Fig. 5, 6, and Pl. III, Fig. 8). These masses project toward the lumen of the tubule and push the spermatozoa toward the center of any cross section picture of the tubule (Pl. II, Fig. 1). The wavy outer margin of the sperm masses is due to the uneven projection of these cell masses into the lumen. In between these masses of rounded, granular cells there appear relatively few cells that are very much larger, less granular, and often are found to be bi-nucleated. In *Rana pipiens* somewhat similar cells have been identified as Sertoli cells and it was clearly shown that each Sertoli nourished at least 28 spermatozoa. Such a relationship is not apparent here. In fact the spermatozoa are arranged quite uniformly in a circle around the lumen of the seminiferous tubule (Pl. II, Fig. 5). Intermediate stages of maturation are not seen. In sexually stimulated males, loose cells are not found among the liberated spermatozoa (Pl. III, Fig. 8) as they usually are in *Rana pipiens.* Such cells have never been seen in the longitudinal collecting tubes, the vasa efferentia, or in the malpighian corpuscles of *Hyla crucifer.*

The spermatozoa are matured before hibernation and are retained in the seminiferous tubules throughout the winter period. However, a very few spermatozoa have been seen in the vasa efferentia, and an occasional spermatozoan has been seen in the renal tubules or Wolffian ducts of hibernating (unstimulated) males of *Hyla crucifer.* As in *Rana pipiens,* and possibly also in man (Wilhelm & Seligmann, 1937) spermatozoa may be liberated in small numbers at all seasons. It must be emphasized that even with the greatest care in manipulation the excision of the urino-genital system may mechanically liberate a few spermatozoa from the testes of untreated frogs, and a false picture of sperm release may result. The tails of the mature spermatozoa may be seen lying freely within the lumen of the seminiferous tubule. It is to be noted that the shape of the spermatozoa of *Hyla* is a spiral as compared with the rod-shaped spermatozoa of *Rana.*

The spermatozoa of *Hyla* (and many *Anura*) are matured in the late Fall, to remain within the seminiferous tubules without further change (Pl. II, Fig. 5) until released during the breeding reactions in the Spring. Temperature and radiant energy may be contributing factors to this release of spermatozoa, but since some Anura do not breed until late Summer or early Fall, there must be other factors at work to control sexual behavior. There is no doubt but that the anterior pituitary gland is the immediate controlling factor, but it has not yet been determined what controls the secretory activity of the anterior pituitary gland of the anuran. The gland does show seasonal variations in potency in respect to sex stimulating factors, and the gonads show seasonal variations in susceptibility to this particular hormone (Rugh, 1937).

The anterior pituitary hormone does not act directly upon the gonads, since excised testes left in the body cavity of pituitary-injected males show no changes attributable to the hormone.

In all probability the hormone is taken into the peritoneal funnels found on the ventral face of the anuran kidney and hence directly to the blood stream—the normal channel of hormone distribution. In these experiments whole anterior pituitaries from adult females of *Rana pipiens* were used.

Within 9–24 hours after the peritoneal injection of the anterior pituitary hormone, gross changes were observed in the testes of *Hyla crucifer*. The solid black hibernating testis changes to a dull gray of the testis with empty seminiferous tubules. Within 9–10 hours after injection (at ordinary laboratory temperatures of 23–25° C.) the spermatozoa are released into the lumen of the tubule (Pl. II, Fig. 6) and within 24 hours many of the tubules have lost the majority of their spermatozoa (Pl. III, Fig. 8). Shortly thereafter only a few scattered gametes may be seen within any of the tubules (Pl. II, Fig. 2).

The spermatozoa are liberated from these oval seminiferous tubules into the longitudinal collecting tubules which are lined with a low type of cuboidal, non-ciliated epithelium. Very possibly there are smooth muscle cells within the testes which (as in the anuran ovary) help to eliminate the gametic products into these exit channels. The motile power for the spermatozoa, after release, may be supplied by themselves at least until they reach the uriniferous tubules where cilia may aid them toward the Wolffian ducts. Three or four of these collecting tubules may be seen in almost any cross section, and may be followed through serial sections until they are seen to emerge as vasa efferentia leading toward the kidney.

THE VASA EFFERENTIA

The vasa efferentia are variable in number and in arrangement even as regards the two sides of the same individual of *Hyla crucifer*. There is generally a network of from 8–12 branching tubes, of at least two different diameters, contributing spermatozoa to the anterior third of the kidney. Some of the tubes end blindly within the mesorchium, but the majority join the kidney on its ventral face near its mesial margin. Each vas branches as it enters the kidney, contributing to several malpighian corpuscles. At most about 40 corpuscles are fed by about 10 vasa efferentia in *Hyla crucifer*.

In the living frog the vasa efferentia can be easily distinguished from the blood vessels and the mesorchium by merely exerting a little pressure on the testis of the hibernating frog. This pressure will force spermatozoa out through these tubes and into the kidneys, causing a flow of whitish cells in contrast with the yellowish stream of large oval erythrocytes in the nearby blood vessels.

Spermatozoa, liberated from the testis by anterior pituitary treatment of the frog, may be seen within the vasa efferentia but their flow is not easily detectable except under high magnification and transmitted light. Spermatozoa can be removed from the hibernating testis by dissection of the gonad in Holtfreter's modification of amphibian Ringer's, or they may be removed from the vasa efferentia, and will prove to be functional as determined by artificial insemination. This indicates that passage of spermatozoa through the kidney is not necessary to the functional maturity of spermatozoa and that the urinary ducts are merely passageways for these gametic products.

The lining of the vas efferens is a low cuboidal type of nonciliated epithelium, very similar to that of the collecting tubes of the testis. At certain regions it appears to be almost a squamous type of epithelium.

The Kidneys

On the ventral face of the kidney, *i.e.*, that portion covered by peritoneum, there are 22–35 peritoneal funnels. These openings are often referred to as nephrostomes but this term is more accurately applied to the opening into a nephridium (Gray, 1930). These peritoneal funnels are most abundant on the midventral face of the kidney and are highly ciliated. As in *Rana pipiens* (Rugh, 1938) these funnels open from the coelomic cavity and their cilia carry coelomic contents into the related funnel tubes and thence into venous sinuses. In *Hyla crucifer* the open mouth of the funnel is only about 10–12 μ across while it is four to seven times larger in *Rana pipiens*. It has been suggested (above) that the presence of these funnels may explain why the injection of hormones into the body cavity results in a quicker response than does subcutaneous implantation, simply because they convey coelomic contents directly into the blood stream. The adrenal tissue, which seems at certain regions to displace as much as 0.1 mm of kidney tissue, is located largely on the median ventral margin of the kidney. It seems excessively large for such a small kidney.

The kidney of *Hyla* averages 3.72 mm in length, 1.16 mm in maximum width, and about 0.5 mm in maximum thickness. It contains from 310–450 malpighian corpuscles in contrast with the 2,000 or more estimated for *Rana pipiens* by Hayman (1928). More than half of the malpighian corpuscles are to be found in the posterior two-fifths of the kidney while the anterior three-fifths includes the 25–40 malpighian corpuscles that convey spermatozoa to the ureter (Wolffian duct) during active breeding. About 10 per cent of all of the malpighian corpuscles found in the *Hyla* kidney are therefore concerned with genital function, in contrast with about 2 per cent in *Rana pipiens*. The majority of malpighian corpuscles are exclusively concerned with uriniferous function.

Relative to the size of the kidney the malpighian corpuscles are tremendous. Those which carry spermatozoa during breeding do not have a maximum size appreciably different from the purely uriniferous corpuscles. Occasionally one sees that portion of Bowman's capsule immediately surrounding the glomerulus literally packed with spermatozoa (Pl. III, Figs. 10 and 12) but it is unlikely that there is any expansion of the walls of the capsule. At certain regions of the kidney the thickness of the organ may be reduced to 0.3 mm, and half of this thickness may well represent a single malpighian corpuscle. Of course the size range of the malpighian corpuscles is considerable, and the size of the related glomeruli is correspondingly quite considerable, the largest of the glomeruli measuring about 0.10 mm in diameter, or one-third the thickness of the kidney.

It is very possible that even during the breeding season the 10 per cent of the malpighian corpuscles which are conveying spermatozoa to the ureters may also maintain some uriniferous function. It is inconceivable that these structures, identical in all respects with the other 90 per cent, should have exclusively a genital function limited to a period of a few days of each year. Spermatozoa have been removed from all parts of the tract, from seminiferous tubules to the seminal vesicle, and they have been found to be functional as tested by artificial insemination. There has been no demonstration of any deleterious effect of urinary secretions on the spermatozoa. In fact, excretory fluids probably prolong the life of spermatozoa by inactivating them in slightly acid medium.

As stated above, the vasa efferentia feed about 40 malpighian corpuscles and these connections are morphologically permanent. These particular corpuscles are located near the mesial margin of the kidney in the anterior three-fifths. The spermatozoa enter the Bowman's capsule at a point opposite the position of the related uriniferous tubule (Pl. III, Figs. 9, 12, 13). The spermatozoa pass around the glomerulus and immediately into the excretory duct. The malpighian corpuscle then is the point at which the urinary and the genital ducts become one and the same. The vas efferens retains its cuboidal epithelium until it merges with the squamous epithelium of Bowman's capsule.

Spermatozoa are found to follow the course of the uriniferous tubule, through its three or four sharp turns (Pl. III, Fig. 9) before it passes dorsalward through the kidney. As it nears the dorsal surface of the kidney it turns at a rather sharp angle and passes toward the lateral margin. As it approaches the lateral edge of the kidney it skirts around the renal portal vein and enters the Wolffian duct, which is found slightly ventral to this blood vessel.

The Wolffian Duct

The Wolffian duct (ureter or mesonephric duct) begins near the anterior tip of the kidney and receives tubules along its entire course. While it starts on the dorsal face of the kidney it gradually takes a lateral position and then leaves this organ at its posterio-ventral tip. For a good part of its course it is found close to the renal portal vein. It is lined with what appears to be transitional (stratified, non-ciliated) epithelium. In any sexually stimulated male frog spermatozoa will be abundantly found within this duct.

There is a slight enlargement and thickening of the walls of the most posterior portion of the Wolffian duct, just before it enters the cloaca. This is designated as the seminal vesicle and particularly in *Hyla* (as in toads) this portion of the urinogenital system is not only a storage place for spermatozoa but is definitely a copulatory accessory. In *Hyla* (as in toads) the eggs are laid singly and must be inseminated singly, hence this lower portion of the Wolffian duct, along with the very muscular cloaca, supplies each egg with a small cloud of active spermatozoa. In *Hyla crucifer,* the male (in amplexus) clamps its cloacal opening just above the cloacal opening of the female at the moment an egg emerges. This process is repeated for each of the 1,000 or more eggs laid, and insures fertilization of each egg. Very probably such body movements as are necessary on the part of the female to eliminate the egg are a signal to the male to liberate spermatozoa. It is quite possible that here, as has been observed in *Rana,* the most posterior uriniferous tubules, which normally have an exclusive urinary function, may receive and temporarily store sperm masses in excess of the capacity of the seminal vesicle. Spermatozoa, to enter these tubules, must do so against ciliary currents.

Summary

1. This paper presents a description of the morphology and the physiology of the reproductive system of the male *Hyla crucifer.*

2. The testis of the hibernating frog is described and the effect of mechanical and of endocrine (pituitary) influences on that testis is demonstrated.

3. By means of pituitary-induced sexual reactions spermatozoa are traced from their seminiferous tubules, through the vasa efferentia, the malpighian corpuscles, the uriniferous tubules, and the ureter to the seminal vesicles.

4. The genital tract, as determined histologically, is very similar to that of *Rana pipiens* where histological evidence was further supported by micro-injection methods. From a urinogenital point of view, *Hyla crucifer* may be considered more like frog than toad.

5. There is no morphological evidence of a Bidder's canal. Spermatozoa enter the uriniferous tract at the point of the malpighian corpuscle. About 10 per cent of the 400 or more malpighian corpuscles of the male *Hyla crucifer* have this dual function.

6. The structure of the kidney is described, including an analysis of the size, number, and distribution of the malpighian corpuscles, and of the peritoneal funnels. The malpighian corpuscles which have a genital function are identified and located within the kidney.

BIBLIOGRAPHY

1. **Beissner**, H. 1898. "Der Bau der samenbleitenden Wege bei *Rana fusca* und *Rana esculenta.*" *Arch. Mikr. Anat.*, Bd. 53.
2. **Bidder**, F. H. 1846. "Vergleichende Anat. und Histol. Untersuchungen uber die Mannlichen Geschlects—und Harnwerkzenge der nachten Amphibien." Dorpat.
3. **Ecker**, A. 1889. *The Anatomy of the Frog.* Trans. by G. Haslam. Clarendon Press, Oxford.
4. **Filatow**, A. 1904. "Entwicklungsgeschichte des Excretionsystems der Amphibien." *Bull. de la Société des Naturalistes de Moscou*, V. 18.
5. Gray, P. 1930. "The Development of the Mesonephros in *Rana temporaria.*" *Quart. J. Micr. Sc.*, **73**: 507.
6. **Hall**, R. W. 1904. "The Development of the Mesonephros and the Mullerian Ducts in Amphibia." *Bull. Mus. Comp. Zool.*, Harvard, **45**: 31.
7. **Hayman**, J. M. 1928. "Notes on the Arrangement of Blood Vessels Within the Frog's Kidney Together with Some Measurements of Blood Pressure in the Renal Portal and Renal Veins." *Am. Jour. Physiol.*, **86**: 331.
8. **Holtfreter**, J. 1931. "Uber die Aufzuchtisolierter teile des Amphibienkeimes." *Arch. f. Ent. Mech. der Org.*, Bd. 124, S. 404.
9. **Hyrtl**, R. 1863. "Uber die Injection der Wirbelthierniere und deren Ergebnisse." *Wiener Acad. Sitzungsb.*, Bd. 47, p. 172.
10. v. **Kolliker**, A. 1875. *Gewebelehre.* Fifth edition.
11. **Marshall**, A. M., and **Bles**, E. J. 1890. "The Development of the Kidneys and Fat Bodies in the Frog." Studies from the laboratories of Owens College, Manchester, V. 11.
12. **Neumann**, E. 1875. "Untersuchungen uber die Entwicklung der Spermatozoiden." *Arch. f. mikr. Anat.*, V. 11, p. 371.
13. **Noble**, G. K. 1931. *Biology of the Amphibia.* McGraw-Hill, N. Y.
14. **Nussbaum**, N. 1886. "Uber den Bau und die Thatigkeit der Drusen." *Arch. f. Mikr. Anat.*, **27**: 442.
15. **Ponse**, K. 1927. "L'evolution de l'organe de Bidder et la sexualite chez le crapaud." *Rev. Suisse Zool.*, **34**: 217.
16. **Rugh**, R. 1935. "Ovulation in the Frog." *Jour. Exp. Zool.*, **71**: 149, 163.
17. **Rugh**, R. 1937. "A Quantitative Analysis of the Pituitary-Ovulation Relation in the Frog, *Rana pipiens.*" *Physiological Zoology*, Vol. 8.
18. **Rugh**, R. 1938. "Structure and Function of the Peritoneal Funnels of the Frog, *Rana pipiens.*" *Proc. Soc. Exp. Biol. and Med.*, **37**: 717.
19. **Rugh**, R. 1939. "The Reproductive Processes of the Male Frog, *Rana pipiens.*" *Jour. Exp. Zool.*, **80**: 81–98.
20. **Spengel**, J. W. 1876. "Das Urogenitalsystem der Amphibien." *Arbeiten aus. d. Zool. Inst. in Wurtzburg*, V. 3, p. 1.
21. **Sweet, Georgiana**. 1908. "The Anatomy of Some Australian Amphibia." *Proc. Roy. Soc. Victoria*, N. S., **20**: 222.
22. **Wiedersheim**, R. 1886. *Lehrbuch der vergleichenden Anatomie der Wirbelthiere.* Jena.
23. **Wilhelm**, S. G., and **Seligmann**, A. W. L. 1937. "Spermatozoa in Urine." *Am. Jour. Surgery*, **35**: 572.

KEY TO PLATES

A—dorsal aorta
B—squamous epithelium of Bowman's capsule
C—Sertoli cells
E—vas efferentia
G—glomerulus
K—kidney
M—malpighian corpuscle
P—peritoneum
R—renal tubules (uriniferous and seminiferous
S—spermatozoa
T—testes
U—seminiferous tubule (testes)
V—posterior vena cava
W—Wolffian duct, ureter, or mesonephric duct
X—renal portal vein
Z—pigment around and within the testes

PLATE I

View of the urino-genital system of the male *Hyla crucifer*, showing the highly pigmented testes.

PLATE I

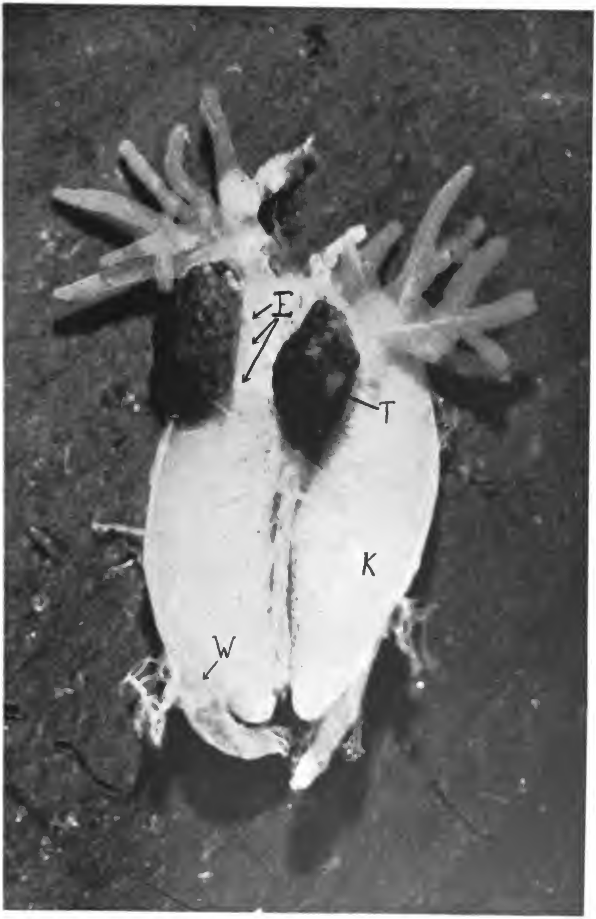


PLATE II

Fig. 1. Cross section of testes of hibernating, or recently hibernating ***Hyla crucifer***, showing mature spermatozoa awaiting release during normal breeding season. Note shortage of interstitial tissue and great abundance of pigment, around and throughout the testes.

Fig. 2. Cross section of testes of frog similar to No. 1 but 24 hours after it had been injected with anterior pituitary hormone, showing many seminiferous tubules emptied of their spermatozoa.

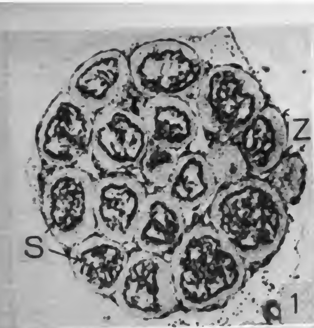
Fig. 3. Cross section of entire urino-genital system of hibernating male ***Hyla crucifer***. Note uniformity of development of spermatozoa in all seminiferous tubules.

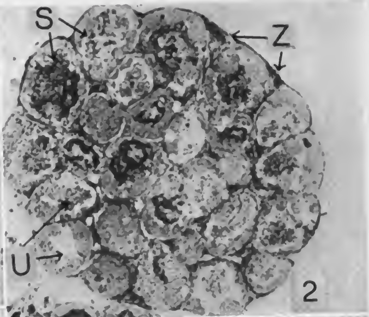
Fig. 4. Cross section of entire urino-genital system of frog similar to No. 3 but 24 hours after it had been injected with anterior pituitary hormone. Note masses of spermatozoa passing through the kidneys.

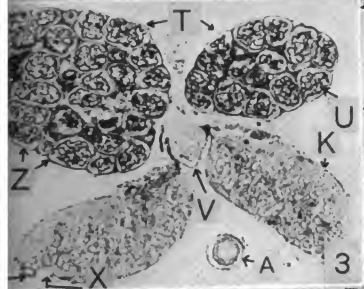
Fig. 5. Greatly enlarged view of a single seminiferous tubule of the testes of an hibernating ***Hyla crucifer***.

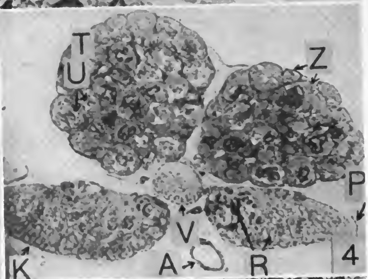
Fig. 6. Similar view of seminiferous tubule of testes from a pituitary-stimulated ***Hyla***.

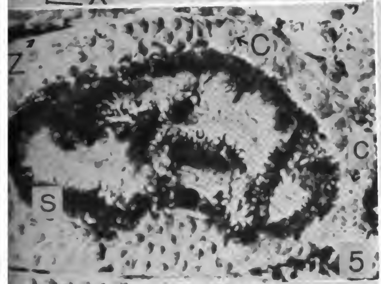
Fig. 7. View of kidney of sexually active male showing spermatozoa within the malpighian corpuscles and the related uriniferous tubules. Also a few spermatozoa show in the ureter.

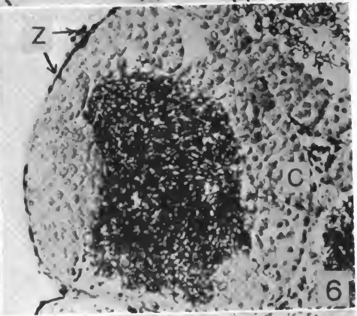

Z
S
1


S
S
Z
Z
U
U
2


T
T
Z
Z
U
K
V
A
X
3


T
U
Z
Z
P
V
R
A
K
4


Z'
C
C
S
5


z
C
6

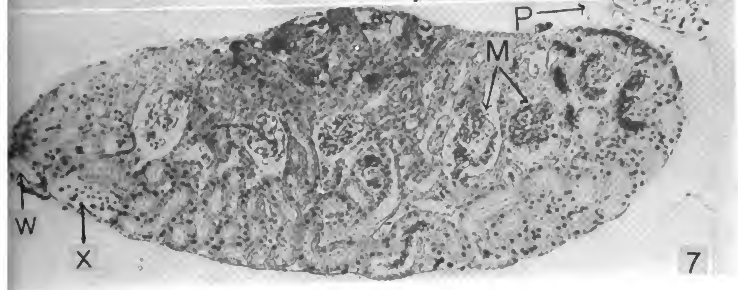

P
M
W
X
7

PLATE III

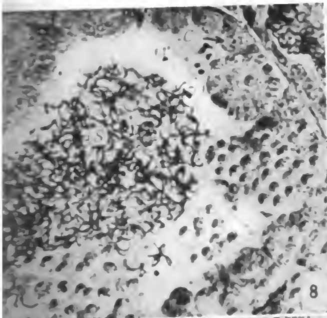
FIG. 8. A single seminiferous tubule 36 hours after the male *Hyla* had been injected with anterior pituitary hormone, showing relatively few remaining spermatozoa.

FIG. 9. Section through inner margin of the kidney of a sexually stimulated male *Hyla* showing (*E*) the entrance point of the vasa efferentia; spermatozoa passing through a malpighian corpuscle around the projecting glomerulus (*G*) and out through the coiled uriniferous tubule (*R*).

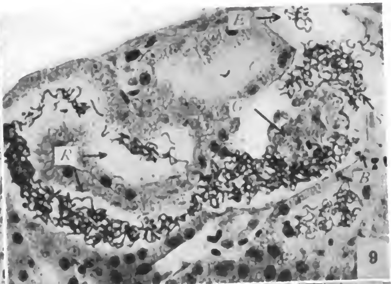
FIGS. 10, 11, 12. Various views of malpighian corpuscles containing spermatozoa.

FIG. 13. Section through kidney showing single malpighian corpuscle containing spermatozoa and many closely related renal tubules filled to capacity with spermatozoa.

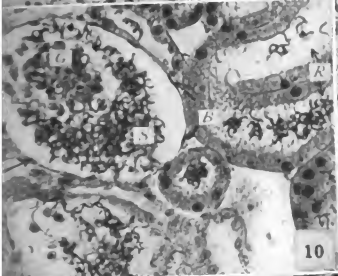
PLATE III



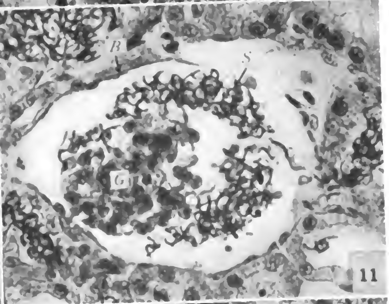
8



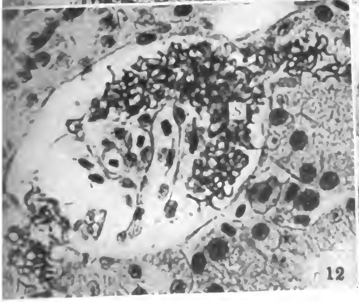
9



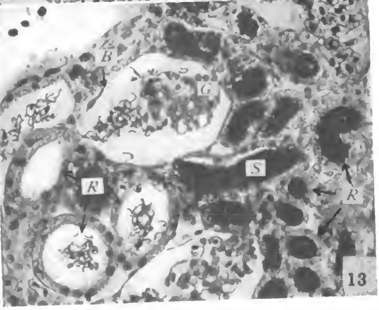
10



11



12



13

WHAT COLUMBUS SAW ON LANDING IN THE WEST INDIES

LEONARDO OLSCHKI

Amherst, Va.

Abstract

The purpose of this essay is to recognize and to explain the reaction of Columbus to the natural and human aspects of the small islands he visited after his first landfall at the Bahamas.

This task suggests a minute and thorough interpretation of the fragments of his *Journal* which contain the account of his first experiences in the New World. The scanty details described by Columbus in these notes obtain a new expressiveness if they are considered as symptoms for his personal or objective interest in the singularities of the newly discovered land.

The investigation of the impulses, emotions and influences which determined the characteristic selection of the things observed and described immediately after this landfall contributes essentially to our understanding of the complex personality of Columbus and to the clarification of the still discussed objective of his enterprise.

In this way it is possible to establish the intellectual attitude of the Admiral toward the reality of exotic life and nature as observed in an unusual geographical environment.

On the other hand it is through this verification of Columbus's reactions and opinions that it is possible to demonstrate the authenticity and reliability of the *Journal* handed down by Bishop Bartolomé de Las Casas.

The question of what Columbus saw on landing in the West Indies may appear an idle one, in view of the fact that we have his letters to Luis de Santángel as well as many fragments of his *Journal,* which contain, together with authentic records of the events of his voyages, the account of his experiences and the direct expression of his opinions and feelings.[1] Thus it might seem that we would be justified in believing that by reading these documents we could grasp the impressions gained by the Admiral during his first contacts with the men and countries of the New World.

But if we give our critical attention to his notes and reports we soon recognize the error in this assumption. For if these were indeed as exhaustive and as reliable as has been supposed, then there would be no "Columbus Question" that becomes more and more intricate as the result of incessantly changing interpretations that teem with more or less fanciful hypotheses.[2]

---

[1] Cfr. Samuel E. Morison, "Texts and Translations of the Journal of Columbus's First Voyage" in the *Hisp. Amer. Hist. Rev.,* XIX, 1939, N. 3.

[2] A bibliographical and critical survey of the present state of the Columbus question is contained in an article of Charles E. Nowel in the *Amer. Hist. Rev.,* XLIV, 1939, N. 3, p. 802 sq.

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 5, JULY, 1941

The insufficiency of the details and the clearly erroneous or inaccurate statements contained in the documents just mentioned reduce the value even of their positive data, and this has given room for doubt as to Columbus's reliability, honesty and good faith.[3] Moreover, from his letters and from the *Journal*, arguments have been drawn that call into question both his geographical knowledge and his nautical capacity—that is, the fundamental conditions of his unprecedented enterprise.

Thus such a distinguished scholar as the late Dr. Cecil Jane sought to demonstrate that at the date of his first voyage Columbus was an illiterate who did not acquire the ability to write until the period between the discovery and the year 1497.[4] If his letter sent to Luis de Santángel in March 1493 does indeed bear in itself, as Dr. Jane assumed, all the marks of having been the work of a clerk who gave literary form to materials supplied to him, then it will be impossible any longer to consider this incomparable document as an exact and personal expression of Columbus's actual opinions.

These far-reaching skeptical conclusions suggest the fitness of a re-examination of the whole question of the literacy of Columbus and of the authenticity of his reports. Recently the expedition organized and led by Professor Samuel E. Morison of Harvard University braved the perils of the ocean and the risks of war, following the track of Columbus's voyages in an effort to check through nautical experience and observation the statements and descriptions handed down by Bishop Bartolomé de Las Casas. Our task is much less pretentious; we simply ask ourselves what Columbus saw in the islands he discovered on his first voyage; we limit our investigations to his first impressions as these are described in the presumably authentic fragments of his *Journal*. In this way we shall not be able to compare his descriptions with the natural and human aspects of the countries he visited. But this restriction of our interest in the documents of Columbus's landfall and discoveries does not necessarily exclude the attainment of some positive conclusions of a wider historical and geographical interest.

---

[3] Especially in the works of H. Vignaud, *Histoire Critique de la Grande Entreprise de C.C.*, 2 vols., Paris, 1911, and R. D. Carbia, *La Nueva Historia del Descubrimiento de America*, Buenos Aires, 1936, intended to prove that the *Journal* of C.C. handed down by Bartolomé de Las Casas is a deliberate falsification.

[4] Cecil Jane, "The Question of the Literacy of C.C." in the *Hisp. Amer. Hist. Rev.*, X, 1930, p. 500 sq. and the Introduction to his edition of *Select Documents Illustrating the Four Voyages of C.C.*, London, Hakluyt Society, I, 1930.

Thus we are led to turn aside from the facts described and to give our attention to Columbus's personal reactions rather than to the objective reality of the nature and peculiarities of the Indies. His impressions are revealed by his words and his style, so that the geographical reality still lives in the form in which he saw it during his astonishing experiences and with the features fixed after repeated observations, and his expressions still have the power to evoke the images of a world long ago vanished and entirely faded from men's memory.

Spanish colonization and the influence of European civilization have thoroughly transformed the natural aspects of the Bahamas and the Antilles that Columbus described. The introduction of a new flora and fauna of European or exotic origin, of field labor, trade and industry, have extensively changed the nature of the soil and the conditions of life in all these islands. There is no trace left of the original Indian natives who were taken away from their native soil and disappeared a few decades after the conquest. Only the larger natural scenery of the mountains, seas, rocks, plains, lakes and rivers, still remains as a stable frame of a human activity which transformed within a short lapse of time a rudimentary stone-age society into a lively colonial organization.

Within these islands very little remains from the age of their first colonization. Furthermore, Columbus's own laconic and summary descriptions of landscapes, localities and events render even more difficult (sometimes indeed impossible) their exact recognition and complicate the attempts to verify his statements by a comparison with the geographical reality and the topographical peculiarities of the islands. It must be borne in mind that the identification of Guanahanì (or, San Salvador) with the Watling Island of the Bahamas was for a long time the result of a compromise reached after inconclusive discussions and accepted in general rather as a provisory and conventional solution than as an established fact.[5]

This uncertainty is due to the circumstance that the terms used by Columbus in describing the island are few and devoid of exactness or color. And it must be added that at this time, and at this stage of nautical science, the exact calculation of the geographical position of newly discovered lands was not within the reach of even the most skillful navigator. All measurements taken in those uncharted waters with the rudimentary instruments then in use could be only approximate and of small practical value.

---

[5] Cfr. the article of R. T. Gould on "The Landfall of C." in the *Geograph. Journal*, London, LXIX, 1927, p. 403 sq.

This fact may explain why the Admiral undertook an astronomical and geodetical determination of his geographic position only in exceptional cases; for the most part he trusted in his own "sens marin" and in the customary tricks of nautical routine. In his *Journal* and in his letters figures are rarely used. The extent of a country, the circumference of an island, the altitude of a mountain are seldom stated in numerical terms, but are generally designated by means of paraphrases, epithets, or apparently insignificant expressions.

Consequently Columbus's descriptions of the lands discovered in all his voyages are to be considered not merely as geographical documents but, at the same time, as literary monuments containing the expression of an individual attitude toward man and nature, science and experience. If we accept this point of view we are protected against every disappointment that might arise from the scientific deficiencies of these reports. At the same time our interest is directed towards the personality of the author, as we try to see with his eyes what the discoveries revealed to him.

After a thirty-six days' voyage from the Canaries, Guanahanì appeared to Columbus at the dawn of October 12th as both a concrete reality and a land of mystery. The small island rose before him as the end, toilsomely attained, of an exciting if uneventful passage and, at the same time, as a starting point toward a land of hope, if not of promise. There he spent two days and two nights, that is, time enough to obtain an exact picture and to gather a comprehensive impression of the country. An unabbreviated portion of the *Journal* containing both the records of these memorable days and the description of the island had been preserved by Bartolomé de Las Casas in a form sufficiently reliable for us to recognize in these fascinating pages the authentic expression of Columbus's first experiences in the New World, and of his reactions thereto.

The greatest part of the description of Guanahanì is contained in the passage written by the Admiral on the evening of the second day (Saturday the 13th of October) in a mood of composure and meditation, when all the natives had gone ashore in their boats.[6] At this very moment he decided to continue his journey in search of the island of Cipango, thereby renouncing all attempt to ascertain whether the cotton and the gold used by the natives were produced on the spot. But in spite of the two days spent at Guanahanì his knowledge of this island was summary and superficial. The (oft-quoted) words devoted to it describe it as "rather large and very flat, with very green trees, many waters, and a large lake in the centre, without any mountain, and the whole land so green that it is a pleasure to look on it."[7]

The only distinguishing feature mentioned in these lines is the lake referred to in the midst of an enumeration of vague and general details. The next day he admired, on the other side of the island to the East, "gardens of the most beautiful trees he ever saw, and with leaves as green as those of Castile in the month of April and May, and much water";[8] but these additional, indistinct, impressions serve to reveal Columbus's feeling for the picturesque aspects of the country rather than to add any new element to the scanty remarks written a day before.

Compared with his more detailed and increasingly enthusiastic descriptions of the islands discovered later among the Bahamas and the Antilles, these first impressions of landscape and vegetation appear even more desultory and conventional. Obviously the Admiral took some time in becoming familiar with the natural peculiarities of the West Indies. These were, on the one hand, strange and baffling to him, but on the other they were not unaccessible to a navigator with some knowledge of exotic countries and climates. Taking up again his navigation, now directed toward the legendary island of Cipango, he seems to have been especially interested in the nautical and strategic aspects of Guanahanì rather than in its landscape and vegetation. For the first time after the discovery he took into consideration the practical benefits of the island calculated from the standpoint of a colonist and exploiter.

---

[6] *Raccolta di Documenti e Studi pubbl. dalla R. Commisione Colombiana,* Parte I, vol. I, p. 18, l.28/29.

[7] *Ibid.*, p. 18, l.15 sq.: "Esta ysla es bien grande, y muy llana, y de árboles muy verdes, y muchas aguas, y una laguna en medio grande, sin ninguna montaña, y toda ella verde, que es plazer de mirarla." In the English translations the word "laguna" is rendered sometimes by "lake," sometimes by "lagoon." It must be kept in mind that, while lagoon denotes in English an area of salt or brackish water (cfr. Murray's *Dictionary*), the Spaniards use "laguna" to designate lake-like stretches of sweet water. Speaking of the Lake Texcoco surrounding the city of Tenochtitlan, the actual Ciudad de Méjico, in his letter to the Emperor Charles V of the 30th of October, 1520, Fernán Cortés emphasized that that capital was "fundada en esta *laguna salada.*"

[8] Cfr. the records of October 14th, *Raccolta* etc., p. 19, l.20 sq.

After having ascertained that the East shore of the small island was surrounded by an extensive reef of rocks making an anchorage impossible and coasting very dangerous, he recognized behind them "a port large enough for as many ships as there are in Christendom", protected by a narrow entrance and by a landtongue which he thought might be fortified and "converted into an island within two days."[9] The extent of this harbour is just as greatly exaggerated as is the number of the islands constituting the archipelago of the Bahamas, which he imagined to be innumerable after the natives who were taken on board his ships "gave him the name of more than a hundred."[10] Likewise he overrated all the distances between the Bahamas visited by him or lying within his sight.

These exaggerations of lengths, quantity, and extent may be explained by the atmospheric conditions of a tropical archipelago that alter the proportions of the objects which are observed and the outlines of the natural scenery. Furthermore, exaggerations of this kind are the consequence of psychological illusions of which Columbus had been the victim in many cases, as had been many other navigators and explorers before him of equal experience. He shared this tendency with Marco Polo, for example, who estimated that the Yang-tse-Kiang carried more shipping than all the waterways of Christendom of his day[11] and that the Eastern "Sea of Chin" in the proximity of Cipango contained 7459 islands frequented by mariners and traders of those parts.[12]

---

[9] *Ibid.*, 1.6–14.

[10] *Ibid.*, 1.25.

[11] *The Book of Ser Marco Polo* trans. and ed. by Col. Yule, 3rd edition, New York, 1929, II, p. 170 and the corresponding passages of the critical edition of L. Foscolo Benedetto, *Il Milione*, Firenze 1928 (Franco-italian version), and of C. Moule and P. Pelliot, *Marco Polo: The description of the World*, 2 Vols. London 1938/39 (Latin version and English translation *variorum*).

[12] Yule, II, p. 264.

Nevertheless, the insufficiency of details, as well as his exaggeration of dimensions and proportions, reveal that the natural aspects of the newly discovered islands did not especially attract the attention of the Admiral. This impression is confirmed by the fact that, on the contrary, he has been very meticulous and exhaustive in giving a very detailed report of the appearance of the natives, their customs and peculiarities, depicting their life and habits with a keen and expressive realism.

Although never considered before as representing a particular problem, the remarkable disproportion between his descriptions of the natural aspects of the island on the one hand, and of its inhabitants on the other, is somehow striking and enigmatical. Against this conventional background, rapidly sketched, of "very green trees, and much water, and fruits of diverse kinds" was described this simple and friendly people, "going naked as when their mothers bore them", represented as "very well made, with handsome bodies, and very good countenances". Again and again, with full details, and sometimes interrupting the account of the events, Columbus mentions the features of these people, the color of their skin, their beautiful eyes and their short, coarse hair "brought down to the eyebrows, except for a few locks behind, which they wear long and never cut". He observes the broad foreheads, the variously painted faces, the straight legs, and the good stature of these "handsome, very well formed people", among whom he noticed no paunchiness, and no physical defects other than the marks of wounds received in fighting foreign invaders. He attributed an especial importance to these witnesses of war and invasion, immediately deducing therefrom that the enemies must "come here from the mainland to take them (the natives) prisoners". This was an entirely unfounded conviction which indirectly reveals the Admiral's opinion about the geographical situation and the political conditions of the islands that were supposed to be in the vicinity of the Asiatic continent and within the reach of the Grand Khan's predatory incursions.[13] These conclusions, drawn from the scars observed on the natives' bodies, are substituted for indications as to the geographical position of the island.

---

[13] On November 1st Columbus noted in his *Journal:* "I believe that all these islands are at war with the Gran Can". *Raccolta* etc. l.c., p. 34, l.9/10.

Columbus was no less impressed by the moral qualities and intellectual capacity of these good-natured savages, and he stresses their harmlessness and docility, noting that, being ignorant of the use of iron, they cut themselves by touching the swords of the Spaniards. On the other hand he was surprised by their vivid receptivity, as well as by the kindness with which they were ready to give away their whole property and their childish happiness in receiving the trash that was distributed among them. His description of the "wonderfully worked" canoes of various kinds and sizes, and of the natives' dexterity in propelling and steering them, betrays the anxious interest with which Columbus observed the lively scenes taking place between the shores and the caravels.

Nevertheless he was not inclined to overrate or to idealize these naive islanders. They soon appeared to him for what they really were—that is, "a race of people very poor in everything" and having no religion, though at his departure they all came to the shore "calling out and giving thanks to God" as a farewell to "the men who had come from Heaven". Later on Columbus corrected these impressions of the natives' religious feelings and conceptions, and he added new details concerning their life and habits, carefully noting the differences between the inhabitants of the other West Indian islands with whom he came into closer and closer contact.

But the impressions gathered at Guanahanì were decisive and lasting. Although there were no more surprises or special events before his landing at Española, he kept alive his interest in the natives, and his curiosity as to their nature; these natives remained for a long time the most important object by far of his penetrating attention. Evidently at Guanahanì and the adjacent islands he was principally concerned with matters of human interest, neglecting the panoramic and natural attractions, and even renouncing his search for the gold mines supposed to be in the vicinity.

If we remember that the exploitation of the natural wealth of the West Indies became soon after an important aim of his voyage, and that the search for gold obsessed him as a fixed idea, we may perhaps suppose that Bartolomé de Las Casas, the apostle and protector of the Indians, extracted from Columbus's *Journal* those passages which especially concern the natives, reducing, in his abstract of the same *Journal,* the records devoted to nautical or natural subjects. In fact, one of the last scholars who sailed in the Admiral's wake for the purpose of establishing the course of the caravels among these islands, suspected that Las Casas failed to include in his précis the astronomical observations of Columbus, probably in the belief that they would not be of interest to the general public.[14]

Such a hypothesis may easily be accepted in our day because of the growing tendency to call into question, not only the reliability, but also the honesty and good faith of this noble, courageous and most certainly conscientious prelate. The trustworthiness of his précis and the authenticity of his quotations are proved, in this case, by the fact that the letter of Columbus to Luis de Santángel resuming his experiences of the first voyage is devoted chiefly to the subject of the natives, and reveals the same lively interest and vigilant care in giving, with more and more details, a complete image of their life and customs. As in the *Journal,* the description of natural scenery takes up only a limited space, while the geographical data appear, for the most part, reduced to the enumeration of names and the—always exaggerated—measures of distances and areas.

As Columbus wrote this letter on February 15th, 1493 "off the Canary Islands" on his return from the West Indies, he already believed that these natives whom he was watching with so much solicitude and benevolence would be sent to Spain as slaves—"as many as their Highnesses (the Catholic sovereigns) shall order to be shipped and who will be from among the idolators".

It is not to be wondered at that for a long time he had had the intention of developing, on a large scale, the slave trade. At that time some of the sailing men were specializing in this kind of traffic, which represented one of the most lucrative branches of the colonial enterprises.[15] The Catholic sovereigns rejected all suggestions to that effect, and the enslavement of the Indians which took place later on was kept within the limits of their own countries.

---

[14] Cfr. R. Cronau, *The Discovery of America and the Landfall of Columbus,* New York, 1921, p. 5.

[15] Gold and slaves were the first cargo brought home by Antão Gonçalves from West Africa in 1441. Cfr. Samuel E. Morison, *Portuguese Voyages to America in the Fifteenth Century,* Cambridge, Mass., 1940, p. 11.

Yet these intentions and actions fail to explain the intense and lasting interest which the Admiral felt for the predestined victims of his undertaking. It is impossible to suppose with any degree of likelihood that his insistent expressions of warm appreciation of the handsome figures and the good nature of these poor savages might have been intended as enticing sales's talks on the part of a slave-dealer anxious to place his wares.

The question of the prevailing human interest of the discovery, apart from its as yet incalculable consequences, becomes even more intricate and important when we take into closer consideration the fact that the reports of Columbus's second voyage deal with the problems pertaining to the first colonial settlement, or with the manners and customs of the Indians, but lack full information about the Lesser Antilles, or about the Admiral's nautical and geographical experiences.[16]

Obviously, we do not possess his own account of the entire voyage, but the *Memorandum* delivered to Antonio de Torres, Captain of the caravel Maria-Galante, as a memorial to the Catholic sovereigns, may be considered as an authentic document written or dictated by the Admiral on January 30, 1494. The most complete account of the whole expedition we owe to Dr. Diego Alvares Chanca, the surgeon of the fleet, whose most valuable information, sent to the Cabildo of Seville in February 1494, is concerned with the Indians of the Lesser Antilles, betraying little interest in the geographical, nautical and natural aspects of the country. Consequently we are even less informed about this voyage than about the first one, although the deficiencies of the two principal documents are made up for by other indirect sources of various origin.

Nevertheless it is difficult to recognize the course taken by that gallant fleet of seventeen sea-worthy vessels, or to identify all the islands mentioned or described in Dr. Chancas's account. His whole attention is directed to their inhabitants. Professor Morison, who recently sought to check the route followed by Columbus between the Lesser Antilles, supposed that the one-sided interest shown by the physician of the fleet for the manners and customs of the Indians was due to the fact that this subject matter was "better calculated than courses and distances, winds and currents, to amuse his municipal correspondents, from whom he expected certain favours and services".[17]

---

[16] Cfr. Samuel E. Morison, *The Second Voyage of C.C.* etc., Oxford, 1939, p. 10.

But it would be difficult to accept the fact that the doctor wrote his very serious and instructive, even if unsatisfactory, narrative mainly for the entertainment of the highest municipal authority of the most important seaport of southern Spain. Nor is it probable that his duties as physician of the fleet prevented him, on the one hand, from following the course of the fleet with the eye of a navigator, and as a scholar, but, on the other hand, allowed him to make precise and detailed ethnological observations. It is more probable that like other navigators, professional or occasional, Dr. Chanca paid no particular attention to the geographical details or the natural and nautical aspects of the country, and that he confined his interests to the life and doings of the natives, as he observed them continuously through direct experience intentionally undertaken.

The predominance of human interest in both the reports of Columbus and those of Dr. Chanca is not a mere coincidence; this is proved by the fact that the same peculiarity is to be observed in the presumably authentic (as well as probably spurious) letters of Amerigo Vespucci. It is not within the scope of our present task to discuss whether the more detailed accounts printed during the life of the latter are forgeries made up from the data of the shorter, and allegedly genuine, letters addressed to Lorenzo di Pierfrancesco de' Medici. But it is an established fact that the great part of the narratives attributed to Vespucci is taken up with more or less extensive accounts of the manners and customs of the natives of the South American mainland and islands, while descriptions of a geographical nature are almost non-existent, and the few that are to be found are still more meagre and perfunctory.

Moreover, it is incontestable that his calculations of geographical position are erroneous or self-contradictory and, at the best, careless and unreliable. His estimates of distances reveal the same tendency toward exaggeration that is recognizable in the accounts of Columbus, while all his remarks concerning the celestial phenomena are sheer nonsense. Nevertheless, all these deficiencies have failed to affect his renown among his contemporaries, or to restrain the learned cosmographer Waldseemüller from glorifying Vespucci in books and maps.

---

[17] *Ibid.*

This conformity of interest to be observed with all navigators of the Columbian era in regard to the aspects of the newly discovered lands goes to prove that the same vivid curiosity linked together the Spanish sovereigns, the members of the Cabildo of Seville, several of the Florentine notables, and the scholars of the Court of King René. In other words, it corresponded with the general tendencies of this epoch, and induced the voyagers to describe the life of the natives with considerable detail, and to neglect the scientific questions connected with the discoveries. Consequently they paid merely a cursory attention to the nautical problems and scientific attainments of their enterprises.

We cannot believe that there was a discrepancy between these general inclinations and the particular interest of the travellers and navigators. It is impossible, or at least improbable, that the latter should have kept to themselves the fruits of their nautical experiences and the results of their astronomical observations, making known only what they supposed to be of wider popular interest. In reality, Columbus, as well as the other navigators of his time, did not see, observe and record in a quite different manner the facts and events which they related, considering these to be essential and the most valuable apart from the practical purpose of their explorations. The time had not yet come when scientific interest was so predominant that the desire for knowledge and understanding would attempt to comprehend the reality of earth and heaven by means of objective measurements and exact calculations alone. These explorers were men whose main intellectual interest, aside from the commercial profit of their enterprises, was concerned with human beings and human peculiarities, just as was the imaginary and symbolic Ulysses of Dante, who crossed the ocean in a spirit of self-sacrifice, desiring

> to be experienced of the world
> and of the vice and virtue of mankind.[18]

It was still the spirit of adventure that prompted and informed the enterprises of Columbus and Vespucci, though of course there were also colonial interests and religious aims involved in such expeditions. Astronomy and cosmography were considered to be useful, but not indispensable, for successful navigation into unknown zones. The resources of these sciences were exhausted in the preparations undertaken for the expeditions. The main objective of the Portuguese astronomers, cartographers, and sailors cooperating in the "Villa do Infante" after 1420, was the realization of the nautical policy of Henri the Navigator, but the attempt which he inspired to establish geographical controls of the expeditions was without scientific method or aim. Astronomy and cosmography still belonged to the "liberal arts" and were cultivated in a speculative spirit, independently of the "mechanical arts" and of the necessities of practical life and specific professions.

---

[18] *Inferno,* XXVI, v. 98/99.

This separation of empirical routine and theoretic speculation, of professional tasks and scientific problems, was characteristic of all arts and disciplines of the Middle Ages and the Renaissance, although successful attempts to overcome the traditional autonomy of both authority and experience were not lacking on either side. The history of medicine, physics and mathematics reveals similar tendencies that represent the typical attitude and the characteristic state of mind prevailing during the epoch of the great discoveries.[19]

Before this period there is no record of a naval expedition that included a cosmographer or mathematician charged with the calculation of geographical positions. Portulans and other maps of navigation were at this time still based on the results of empirical routine, and on nautical traditions which had nothing in common with erudition or science.[20] It was long after the discovery and colonization of the West-African coast that Master Joseph, a physician and cosmographer, with a nautical astrolabe of new construction, executed in Guinea several valuable astronomical observations which were mentioned with appreciation by Columbus in his autographical notes to Piccolomini's celebrated *Historia*.[21] But this same Joseph, together with his colleague Roderigo and Diego Ortiz, Bishop of Ceuta, was responsible for the rejection of Columbus's first official project of an oceanic navigation, which in 1482 he submitted to King Joãn II of Portugal.[22] In the course of the following ten years the famous "junta" of Salamanca and a *primariorum hominum consilium,* composed of Spanish theologians, scholars and sailors, refused to accept his arguments in favor of his enterprise.[23]

---

[19] These aspects of the History of Science of the Middle Ages and the Renaissaince have been considered in the writer's *Geschichte der Neusprachlichen Wissenschaftlichen Literatur*, especially in Vol. I, Heidelberg, 1919, and II. Leipzig, etc. 1922.

[20] Cfr. the author's *Storia Letteraria delle Scoperte Geografiche*, Firenze, 1937, p. 143 sq.

[21] "Rex Portugalie misit in Guinea A.D. 1485 magister Jhosepius, fixicus eius et astrologus, [com] piendum altitudinem solis in tota Guinea; qui omnia adimplevit et

Its successful realization did not depend on the assistance of scholars and men of science; in the long years of struggle and preparation Columbus found among intelligent and cultivated men of practical experience the understanding and the support which the representatives of official science had denied him. Thus, returning to Spain, after his third voyage, as a prisoner of Bobadilla, the Admiral haughtily refused to be judged by the "Caballeros de Letras"—*i.e.,* those suspect knight of the quill that had always denied him their aid.[24]

Active cooperation between navigators and cosmographers began shortly after these successful oceanic enterprises in order to give a systematic basis to the nautical art. The correspondence between Columbus and Toscanelli (if this be authentic) anticipated this development to some extent, but in any case the Genoese sailor did not expect, nor did he receive, from the Florentine scholar, anything more than statements and deductions of a purely speculative character.[25] He believed these to be sufficient to convince the sceptics of the Portuguese and Spanish courts with the force of the theoretical arguments contained therein; as is well known, the results failed to match his expectations.

It is now understandable that, after such experiences, Columbus engaged no scholars, scientists, or priests to join his first expedition; all his companions and subordinates were sail-

---

renunciavit dito serenissimo regi, me presente etc., et hoc cum maxima diligentia procuravit.'' *Raccolta* etc. P.I, Vol. 2, p. 369, § 860. Master Joseph is mentioned again in the marginal notes to Peter d'Ailly's *Imago Mundi* in connection with the measurement of the degree. *Raccolta* etc. l.c., p. 407, § 490.

[22] For more details cfr. Cesare de Lollis, *C.C. nella Leggenda e nella Storia,* 3th ed., Roma (1923), p. 73 sq.

[23] *Ibid.,* p. 81 sq. and 95 sq.; J. B. Thacher, C.C., I, 1903, p. 417 sq.

[24] Letter to the Nurse of Prince Juan. *Raccolta* etc., P. I, Vol. 2, p. 73, l. 6; Thacher, *op. cit.,* II, p. 423.

[25] The text of the Letters in *Raccolta* etc., P. I, Vol. 1, p. 364 sq.; Thacher, I, p. 300 sq.

ors and routine men of different practical professions.[26] And, probably for the same reason, he refused to engage a competent astronomer as navigating officer for his second voyage and, at the end of his career, expressed the belief that he was directed in his discoveries by prophetic visions and by the will of God. This certainly was not his opinion during the period of his supposed correspondence with Toscanelli, when he was seeking to gather from various sources scientific evidences as an aid for his projected voyage.[27] Nor could it have been his belief to be "a man sent from God" that determined him to refuse the mathematician suggested by the Queen; undoubtedly he was convinced that the assistance of this learned astronomer would have been devoid of practical benefit. As the objective of his enterprise was to reach an (undisclosed) section of Eastern and Southern Asia vaguely designated as Indian and described by classical authors as well as by Marco Polo, Mandeville, and Toscanelli, it is probable that Columbus had in mind the remarks of Solinus concerning the navigation in those waters, where, as the Latin geographer affirms, observation of the stars was useless or even impossible.[28] Through Peter d'Ailly's *Imago Mundi* he had learned of this passage of Solinus,[29] whose statements of a descriptive and imaginary character were propagated in all the works of scholarly and popular geography of the period.[30] In any case, neither Columbus nor Vespucci—a less pious Florentine and a man inclined to an even more realistic evaluation of such expeditions—engaged cosmographers or mathematicians to assist them in their navigation. Evidently the confidence of these men in the resources of science was limited, being confined to the theoretical premises of their undertakings. They were adventurers who trusted in their experience, intelligence, and endurance, expecting to find wonders and adventures—and profits—in unknown lands.

---

[26] The names of Columbus's companions and of the members of his crew are mentioned—in Alicia B. Gould y Quincy's "Nueva Lista Documentada de los Tripulantes de Colón en 1492" in *Boletin de la R. Academia de la Historia*, Madrid, 1924–1928.

[27] The late Dr. Cecil Jane insisted on the assertion that Columbus considered his enterprise, since the beginning, as a divine mission directly inspired by God. Cfr. the Introduction to *Select Documents* etc.

[28] "Nulla in navigando sidera observatio, utpote ubi septentriones nequaquam videntur vergiliaeque numquam apparent". (C. Julii Solini, *Collectanea rerum memorabilium*, ed. Th. Mommsen, Berlin 1864, p. 218.) "Observatione itaque navigandi nulla suppetente, ut ad destinatum pergentes locum capiant, vehunt alites etc." (*ibid.*).

[29] E. Buron, *Ymago Mundi de Pierre d'Ailly*, Paris 1930, II, p. 395. The passage was partially underlined and annotated by Columbus (*Raccolta* etc., P. I, II, p. 395, § 320).

[30] Cfr. Ch. V. Langlois, *La Connaissance de la Nature et du Monde au Moyen Age*, Paris 1927. Mandeville (*Travels*, Chapt. XXXIII) affirms that "in those islands (*i.e.* of Taprobane and the East Indies) men see there no stars so clearly as in other places. For there appear no stars, but only one clear star that men clepe Canopus".

The reports of all mediaeval travellers show an identical standpoint as regards the selection and description of the facts that were taken into closest consideration. Their interests were directed mainly to earth and man, occasionally to beasts and monsters, sometimes to supposed wonder-working enchantments. The increase of realism in the accounts of the travellers after the Franciscan missions in Asia during the 13th century did not affect this general tendency. In reports of this type, as well as in all fiction and poetry of the Middle Ages and the Renaissance, Nature constitutes an accessory element of the narrative, being merely a frame for human life, activity, and events. For example in the romance of *Amadis,* where the principal events take place in imaginary islands of phantastic archipelagos, landscape and natural scenery appear exclusively in this function. And they are represented with no more words and in no different terms than those which Columbus, at the same time, used to describe the natural aspects of the Bahamas, observed by him immediately after his landing at Guanahanì. It seems that, at the beginning, he gave no great attention to them, so that these first descriptions are lacking in color and detail.

This attitude toward Nature and the natural landscape was characteristic not only of all the explorers of Africa and America during the 15th century, but also of Marco Polo and the travellers before and after him. The essential difference between the oceanic navigators and these continental travellers lies in the fact that the latter came into contact with highly civilized peoples or with nations influenced by such peoples, while the others discovered primitive or savage tribes who lived in a state of nature.

If we compare, for instance, the account of Aloise da Cadamosto's voyages to the western coast of Africa with Columbus's *Journal* and Letter of the first voyage, or with Dr. Chanca's report, we shall find a similar state of mind revealed, according to which the purport and the substance of their records are limited to a few arguments of an analogous nature.[31] On his voyage from Cape Saint Vincent in March 1455 to the Canaries and thence along the African coast as far as Senegal, the Gambia, and the Cape Verde Islands, the young Venetian nobleman and trader was particularly interested in the inhabitants. The emphasis of his narrative is placed on the description of their stature, appearance, life, customs, houses, markets, products, and beliefs; this is followed by a vivid and detailed enumeration of the strange animals living in that region. But the geographical details are limited to cursory remarks and to the names of countries, islands, capes and rivers; the whole is just sufficient for a rapid orientation. Nor is there a word concerning the change of the stars or the aspect of the sky.

A century before, Giovanni Boccaccio had given, in the same style, a colorful description of the natives of the Canary Islands, availing himself of the information which had reached Florence from Seville and which concerned the expedition organized in 1341 by Niccoloso da Recco and Angiolino de' Corbizzi.[32]

Taking a broad view of the evolution of travellers' accounts, as representing a particular branch of the literature of the late Middle Ages, we may infer from all these coincidences that the records of Columbus do not deviate substantially from a tradition which had been kept alive by a lasting spiritual attitude and which corresponded with the culture, the education and the interests of all prominent voyagers of the epoch of the great discoveries.

It is possible to draw some important conclusions from this statement, but first we must abandon the general opinion which presupposes the existence of a certain discrepancy between the voyagers as authors and the public as readers. The *Journal* of Columbus was not composed with the intention of impressing the sovereigns and the court with tempting prospects of further expeditions even more successful; on the contrary, it must be understood as representing a series of monologues of a vivid and intense spirit approaching a new field of experience and disclosing his impressions and emotions in frank words, genuine and unadorned. Thus, from time to time, there is in his descriptions and pictures a lyrical note arising, on especial occasions, to higher accents, in which a sympathetic sensitivity may recognize a poetical vein and an aspiration toward a nobler and more artistic style.[33]

---

[31] R. Caddeo, *Le Navigazioni Atlantiche di A. da Cà da Mosto* etc., 2nd edit., Milano 1929. English translation with Introduction and notes by G. R. Crone, *The Voyages of Cadamosto*, London, Hakluyt Society, 1937.

[32] Caddeo, *op. cit.*, p. 141 sq.

This unmistakable ring of Columbus's prose, as well as the bulk of the positive remarks contained in the *Journal*, prove the genuineness and the reliability of Las Casas's quotations and abstracts. Nothing essential has been left out in the latter's précis—at least, in the passages concerning Columbus's landfall and his navigation among the Bahama Islands, nor did the original contain more details of his navigation than those revealed by the unabbreviated or the integral passages of the *Journal*.

It was not until his third voyage that Columbus, sailing near the equinoctial, undertook a more systematic observation of the position of the stars, each night "marvelling at such a change in the heavens".[34] At this point Las Casas refered to the Admiral's astronomic observations as carried out, it would seem, in a poetical mood rather than with scientific procedure. This does him more credit than if he had taken calculations, and made observations, which would inevitably have turned out to be erroneous and misleading. For the rest, he was a man of his time, accustomed to consider the world "sub specie humanitatis" in spite of his religious piety and exalted devotion. He followed the anthropocentric tendencies of his epoch, which transformed the earth into a stage displaying a gay tableau of various human singularities.

While the geography of scholarly tradition still peopled the exotic regions of the earth with all sorts of imaginary monsters, Columbus looked upon this "handsomely formed people" of the Lucayos with the delight and the appreciation of an artist of the Renaissance. Vespucci did the same a few years later, expressing his admiration for the well proportioned bodies of the South American natives, whose barbarous customs he described with keen interest but less benevolence.

---

[33] Alexander von Humboldt pointed out repeatedly, and with admiration, the peculiarities of Columbus' style, generally disregarded by the Admiral's biographers. (Cfr. *Examen Critique de l'Histoire de la Géographie du Nouveau Continent*, passim, and especially Vol. III, Paris 1837, p. 227 sq. and *Cosmos*, II, A, Chap. 1.)

[34] *Raccolta* etc. P. I, Vol. II, p. 24, l. 10 (Las Casas précis of Columbus' narrative of the third voyage).

During his cruise from island to island Columbus's attitude seems to have changed very slowly. The adjustment of his mind to the unusual aspects of the country was generally cautious and groping, and he was seldom dazzled by his enthusiasm, in spite of the exaggerated expectations which he held to tenaciously during the whole way.

If we consider the notes of the *Journal* in themselves, avoiding the common tendency of his biographers to make an epic or a dramatic paraphrase out of his own unadorned and dispassionate expressions, then we may recognize, through Columbus's words, what actually went on in his mind, influencing his decisions and inspiring his opinions. In this way we become able to fix the limits of his imagination and to measure the extent of his sense of realism.

Sailing between the Bahamas or along the shores of Cuba and Haiti, Columbus steadily developed his ability of distinguishing the different aspects of their landscapes and exotic vegetation, as each newly discovered land seemed to him more beautiful than the one last described. It is an extremely interesting task to follow the different stages of this spiritual conquest of the New World as it progressed, day by day, in the Admiral's consciousness. This, however, should be the subject of a special investigation; at present we are concerned only with his first impressions, considered as the starting point of his gradually widening interest in the islands and as a reliable measure for the increasing dimensions of his intellectual horizons. His concluding remarks about the landscape and vegetation of the Bahamas prove that his discernment and penetration in experiencing and observing the New World failed fully to reveal themselves in sight of its natural aspects.

On October 15 Columbus landed at Santa Maria de la Concepción, which is supposed to be the Rum Cay of the Bahamas; he went through the island paying attention mainly to the actions of the natives as observed in his intercourse with them. As for the landscape he notes only that "these islands are very green and fertile, and the breezes very soft", deducing (on more than one occasion) from this fact that they might be hiding many valuable but unknown things.[35]

---

[35] *Ibid.*, P. I, Vol. I, p. 21, l.16, The passage reads in the Spanish text, "ayres muy dulces".

The description of the island points out no distinctive traits, and the terms employed seem to be even poorer and emptier than those used a few days earlier in reference to the shores and the "gardens" of Guanahanì. But after the statements and comments above, it would be an error of historic perspective to ask for more objective details and personal impressions, or to infer from the deficiencies of the descriptions that Columbus was indifferent to the charms of the West Indian landscape.[36] In reality we can easily establish that, in the further course of his explorations, his descriptions became more and more eloquent, circumstantial and lyrical, as the natural aspects of the islands visited by him seemed more and more to be in harmony with ideal landscapes of poetical features, animated by lofty mountains, gleaming rivers, evergreen trees, and singing nightingales.[37]

But obviously, though considered as belonging to "India", the two islets first discovered in the Bahamas seemed still to be far from Ophir, Paradise, and the enchanted wonderlands of the East. To Columbus these small, flat, green islands were pleasing and curious indeed, but evidently not as attractive or tempting as the larger Bahamas, Cuba or Haiti appeared to him later. And in the course of his cruise Columbus's appreciation was quite justified, revealing by this graduation of impressions and expressions the sincerity of his emotion and the honesty of his words.

In spite of the pleasure he took in looking upon them, these islands inhabited by "a people poor in everything", were not sufficiently valuable or promising to detain the ambitious Admiral and his escort of unsentimental sailors and adventurers. Therefore he left to go about seeking traces of the things he was expecting to find in the discovered region. The vegetation, though consisting "of the most beautiful trees he ever saw", seemed to promise little as to the value of their produce. The small bundles of cotton offered by the natives hardly represented precious merchandise, nor could the tiny pieces of gold which he saw dangling from the noses of some of the natives have given him proof of the supposed vicinity of Cipango.

---

[36] Cfr. Cesare de Lollis, *op. cit.*, p. 138, and R. Menéndez Pidal's remarkable article on "La Lengua de Cristóbal Colón" in *Bulletin Hispanique*, Bordeaux, 1940, p. 1 sq.

[37] Cfr. the *Storia Letteraria delle Scoperte Geografiche* quoted above (n. 20), p. 11 sq.

Consequently, what Columbus saw on landing in the West Indies, though perhaps sufficient to confirm a preconceived opinion about the exotic countries he expected to discover, could, by itself, certainly not have been able to inspire the visions which he had of the riches and wonders of Eastern Asia. This proves the groundlessness of Vignaud's assumption that it was only when he was in sight of the Bahamas that Columbus decided to direct his expedition toward Cipango and the empire of the Grand Khan. No less untenable, from this angle, appears the opinion of Dr. Jane, that the great enterprise of crossing the ocean was without a definite geographical objective.[38]

From the first impressions gathered soon after his landfall, we are able to recognize the circumstances which induced Columbus to believe that he had reached the borderlands of the Asiatic continent for which he had set sail more than three months before. Beholding the seemingly innumerable islands of the Bahama group, he certainly identified them as belonging to the huge insular region so impressively described by Marco Polo and indicated in the map which the Admiral carried on board.[39] Besides the general impression he received of the fertility of the soil, the luxuriant vegetation, the apparently evergreen trees, and the "very soft breezes" seemed to correspond with the widespread image of India and the Far East, as sketched by all the ancient and mediaeval authors dealing with geographical matters familiar to Columbus, and irrefutably confirmed by Marco Polo's enthusiastic descriptions.[40] These were the inducements which stimulated the Admiral to start immediately his search for gold, even though the condition of the natives seemed not very encouraging for this purpose. His confidence of success came from the expectations which he and his companions brought with them from Spain. These were all men who had grown up and been educated in an epoch when logical deductions drawn from supposedly indisputable premises still were, at the least, as conclusive as practical experiences.

---

[38] Cfr. the Introduction to his edition of *Select Documents* etc. and to the English translation of *The Voyages of Christopher Columbus*, London, The Argonaut Press, 1930).

[39] It was only on November 14th that Columbus expressed the belief that he discovered "those innumerable islands that are depicted on the maps of the world in the Far East" (*Raccolta* etc., P. I, Vol. I, p. 42), but the mention of Cipango immediately after the discovery of the Bahamas (*ibid.*, p. 18) proves that he certainly included the latter in the same region of Eastern Asia represented in the maps of Fra Mauro and Martin Behaim, and in that of Toscanelli supposed to have been on board of Columbus's caravel.

[40] Cfr. Peter d'Ailly's *Imago Mundi*, ed. E. Buron, Vol. I, Ch. 15 and 16, p. 258 sq.; Mandeville's *Travels*, Ch. XXXII sq. and the treatises condensed and annotated by Ch. V. Langlois, *op. cit.*, passim.

As the natural aspects and the lovely climate of the country seemed to confirm their belief that they had landed in a country belonging to "India", the existence of gold somewhere in the immediate geographical neighborhood was assured for them all by simple and uncontestable implication, in spite of the fact that the traces of the precious metal were insignificant, and the gold mines themselves invisible.

The discrepancy between Columbus's increasing confidence in his excited search for gold, and the disappointing lack of signs indicating its actual existence in the newly discovered islands, has given rise to unfounded and misleading speculations which have deeply affected the opinion of his character and personality and which, at the same time, have altered and entangled the essential historical problems connected with his achievements. It is out of these assumptions that the doubts as to his reliability and good faith have grown up in recent literature devoted to the Admiral, who is represented in biographies, fiction and essays as a fraudulent adventurer, a greedy pirate, an exalted fool, a professional impostor or, in the best of cases, as an anticipated Don Quixote of the Ocean.

Placing his emphasis on this striking contradiction between actual experience and wishful interpretation of facts, even the best informed of the Italian biographers of Columbus overlooked, minimized, or denied all the other interests of the discoverer in the natural aspects of the islands;[41] consequently he considered Columbus's descriptions of landscape merely as rhetorical padding or as irrelevant and conventional utterances. In a book as eloquent as it is unconvincing, Salvador de Madariaga recently expressed the belief that the intense and tenacious greediness of Columbus for gold represented a typical manifestation of the Jewish character; in accordance with this ancient prejudice he concluded that the Admiral was the scion of a Spanish family of *conversos* who had emigrated to Genoa in the fourteenth century.[42]

---

[41] Cesare de Lollis, *op. cit.*, p. 12 sq. and 138.

All these authors ignore the fact that a thirst for gold and similar visions, more or less ecstatic, of wealth and lustre, constituted, in the age of the discoveries, the principal economic inducement that inspired the Portuguese navigations to the West African coast and India, the Spanish conquest of the New World and, a century later, the expedition of Sir Walter Raleigh to Guiana. Columbus started from Spain with just such visions and expectations in his mind, and no premature disappointment was able to shake his confidence or the power of his arguments based on old traditions and authoritative premises. Gold represented the only profit immediately realizable of such costly enterprises, and provided the most direct means of financing an oceanic expedition in this critical epoch of Spain's economic life. This is the simple and the true reason for the predominance of the gold-motif in Columbus's *Journal*, quite apart from pretended racial influences or personal instincts.

It is a matter of fact that, shortly after landing at Guanahanì, he "was attentive and took trouble to ascertain if there was gold",[43] and that the search for it became even more intensive during the course of the cruise. In all places, his eyes were directed toward every thing which might have the glittering appearance of gold. We may admit that, at the period of the great discoveries, this was the common attitude of all the explorers and conquerors in search of riches in three different continents. But before Columbus none of them wrote a *Journal* revealing, day by day, their reactions and expectations. His records prove that, from the time of his departure from Palos, Columbus had been certain of his geographical goal, and aware of the incomparable importance of his expedition, directed toward a country partly real, partly fabulous. Consequently, the interplay between experience and imagination which is revealed in his records and observations, was determined by his expectations as well as by his vivid and enthusiastic spirit. He was not the only victim of such visions and phantoms. Fascinated by the records of the *Journal* his biographers have too often overlooked details which were more consistent and more valuable, preferring rather to hunt out his alluring and misleading illusions than to consider the progress of his personal experiences. But every cautious reader of these records is able to state that the horizon of Columbus's interests increased simultaneously with his sense of realism, in proportion as the gold mines of Cipango and Babeque vanished in the background of the amazing and troubling West Indian scenery.

---

[42] The other allegedly Jewish traits discovered by Madariaga in Columbus's character are his "bargaining sense" and a "typically Jewish mobility". Apart from the fact that the pretended Jewish origin of Columbus could never explain his personality and his achievements, one must deeply deplore that such perfunctory commonplaces, used by unscrupulous agitators for the purpose of religious and racial persecution, are employed now by talented and non-sectarian authors as a standard for the interpretation of historical personalities and events. Cfr. Samuel E. Morison's book review of Madariaga's biography of Columbus (New York, 1940) in the *American Historical Review*, XLV, 1940, N. 3, p. 653 sq. A few lines in R. Menéndez Pidal's article quoted above (n. 36) are sufficient to undo Madariaga's speculations about the supposed Jewish origin of the Admiral.

[43] October 13th, *Raccolta*, etc., P. I, V. I, p. 18, l. 6.

The different attitudes of the Admiral in regard to empirical reality and remote possibilities, are already recognizable a few days after his first landing. His principal reactions may be observed in the direct and simple expression of his opinions and in the description of incidental events. For example, as he anchored, on October 15th, at Santa Maria de la Concepción, a few miles distant from Guanahanì, he wanted to ascertain if the island contained gold—for he had been told by the natives taken on board that "there they wore large bracelets on the legs and arms".[44]

After some doubts, at first, as to the veracity of such reports, he continued to entertain the idea of people going laden with gold and jewels, an idea which confirmed and enhanced his hope and confidence in the success of the expedition.[45] It can be taken for granted that this enticing picture corresponds rather with popular ideas of Oriental wealth and luxury prevailing in Mediterranean tales, than with the imagination of the simple West Indian natives. Nor is there need to remind ourselves that neither Columbus nor any other explorer of America was able to find, in any place, people wearing such kinds of ornaments. The Admiral who interpreted in this manner the unintelligible talk of his Indians was the victim of the same psychological illusions that lead us to hear sweet melodies in the chime of church-bells, or to discover in the clouds familiar features or impressive images of phantastic shapes.

---

[44] Oct. 15th, *ibid.*, p. 20, l.5 sq. and 21, l.5 sq., 18 sq.

[45] A month later (12th of November) the Admiral was quite convinced "that in these islands there are places where they (*i.e.* the Indians) dig out gold, and wear it on their necks, ears, arms, and legs, the rings being very large."

In dealing with his search for Cipango and with all the phantoms which lured him, the biographers of Columbus have generally emphasized, and indeed excessively, the power of these deceptive images and the consequences of his misinterpretation of reality. But it is no less significant for his character, and no less interesting for the history of his enterprise, that such winged illusions did not hinder him from observing and describing, though with apparently insignificant details, the impressive facts and events of his discoveries. We have eloquent proof of this assertion shortly after his arrival in the West Indies.

It must strike every attentive reader of his *Journal* that, after having consistently evoked the phantastic image of a people covered with gold from head to foot, he immediately begins to give an impressive graphic description of an Indian coming "alone in a boat on his way from the island of Santa Maria to that of Fernandina, carrying with him a piece of their bread, about as large as the fist, and a gourd of water and a piece of brown earth, powdered and kneaded".[46] Columbus also noticed "some dried leaves—probably tobacco—which must be a thing highly prized among them"; this he had already observed at Guanahanì. But the only objects of value owned by the man were a string of beads and two small Spanish silver coins worth less than a dime each and "kept in a basket of their making".[47] This poor treasure of the lonely savage was taken along with him as evidence, as he went in his boat from island to island announcing the arrival of "the men who have come from heaven". A scene no less pathetic and impressive than are some of the more glorious episodes of Columbus' adventurous life.

There is an affecting contrast between his visionary expectations and the unadorned reality which he describes with such symptomatic details. A small piece of gold hanging from a hole which some of the natives had in the nose was for him, on the first day of his stay in the West Indies, indisputable proof of the existence in that country of the precious metal. And on the other hand, the entire equipment of the Indian errant in his boat was sufficient to reveal to the Catholic sovereigns and to every reader of the *Journal* the poverty and frugality of these new subjects of the Spanish crown.

---

[46] October 15th, *Raccolta* etc., *loc. cit.*, p. 21, 1.20–30.

[47] *Ibid.*

Thus we see that it was far from the Admiral's intention to mislead his sovereigns by creating false impressions; we must recognize that he composed his *Journal* in a spirit of perfect sincerity. Wishful interpretation of the objects to which his attention was attracted may have stimulated his enthusiasm and his energy, but this was without self-deceit, and without any intention to fool his companions, the rulers of Spain, or the public. Against the golden background of the marvels of Cipango this precise observation of the smallest details of an apparently trivial incident of his voyage is exceedingly eloquent and significant. We may deduce from this first instance that his mind embraced, with an energy of effective comprehension, the suggestions of imagination as well as the appearances of reality. Sometimes these two tendencies may have interfered and led him to false interpretations and erroneous judgments. But without this power of imagination he never would have been able to discover something in the reality of exotic life and nature.

These tendencies were correlative, complementary, and not opposed the one to the other as so many biographers have represented them, thereby creating, and inflating with ingenious dialectical artifices, dramatic contrasts. It was this permanent intentness of his intellectual power that transformed the optical perceptions into exact, purposive observations, the range of which gradually extended. He looked at reality with a spiritual eye and with a breadth of interest and a multifariousness of impressions that has no parallel in the history of voyages and discoveries. And it is in consequence of this wide extension of his impressionable spirit that Columbus was a great man and not alone an able navigator and explorer.

Likewise in Galileo's life and achievements there have to be considered not only his exciting discoveries in the heavens, and his dramatic, if not scientifically unobjectionable, defense of the Copernican system; conversely, his stupendous observations and speculations concerning apparently more trivial phenomena (*e.g.* the fall of bodies, the oscillations of the pendulum) should be valued and studied with no less admiration—both in themselves and as the starting points of a new spiritual evolution of science and mankind.

It is time now to detach Columbus's figure from the golden background against which he is wont to be painted, and to consider him in the midst of the reality he described, described with personal accents, it is true, but none the less, with an increasing insight into the peculiarities of life and nature in the islands he visited. Thereby he became an explorer who extended his discoveries to new spheres of human experience and knowledge. The illusions that possessed him were founded on an authority which owed itself to the unshakable mediaeval belief in religious dogmatism and bookish traditions. The lack of scientific details, the erroneous statements of a geographical or astronomical nature, and his neglect of the aspects of life and nature represent the limits of his receptivity and of the intellectual and cultural horizons of his time. But his independent approach to the natural aspects, the landscape, the vegetation, the climate, the birds and animals of the West Indies, is personal and original, on the one side influenced by a poetical temper, on the other determined by an uncommon power of realistic observation.

His descriptions of men and beasts, of trees and rocks, of mountains and harbors, of forests and rivers, must be considered in connection with the entire evolution of travel narratives and descriptive literature. In this way all the apparently unconnected and necessarily imperfect observations undertaken by the Admiral after his landing in the West Indies appear as slow and circumspect steps directed toward the spiritual conquest of the New World. Thus, even the small and incidental descriptive details concerning the natural aspects of the islands, which are contained in the *Journal* and the Letters of Columbus, gain an increased significance, and may claim a thorough and minute investigation of their intrinsic and comparative value.

For such an investigation, only the fragments of his *Journal* possess a documentary interest. The indisputable authenticity of the passages devoted to his landing in the West Indies furnishes the criterion for the reliability of Las Casas' *précis* and quotations. Even in this form the *Journal* serves to reveal to us that this New World disclosed itself to the eyes of Columbus little by little, as he learned to distinguish, in his emotional and intellectual reactions, the characteristic marks of this unfamiliar environment. And, therefore, what he saw and what he failed to see on landing in the West Indies is equally significant for an understanding of the Admiral's personality and of the results of his fateful enterprise.

THE EFFECTS OF HEAT AND ULTRAVIOLET LIGHT ON CERTAIN PHYSIOLOGICAL PROPERTIES OF YEAST

THOMAS F. ANDERSON

Formerly Research Associate in Plant Physiology, University of Wisconsin

and B. M. DUGGAR

Professor of Plant Physiology, University of Wisconsin

ABSTRACT

The effects of irradiation and heat on yeast were studied with particular emphasis on changes in physiological properties which could be observed during or immediately following treatment. Wave lengths shorter than 2200 Å were found to reduce immediately the rate of respiration and cause cells to stain with methylene blue. Irradiation with λ2650 caused little reduction in the rate of respiration, but prevented the normal increase of respiration in nutrient. Exposure to a temperature of 50° C. rapidly reduces the rate of respiration, but brings about only a gradually increasing tendency of the cells to stain with methylene blue. All these treatments rapidly reduce the colony-forming ability of the cells.

When heat and radiation treatments were combined it was found that irradiation sensitized both the colony-forming ability and the staining resistance to subsequent heat treatment. On the other hand, the rate of respiration was found to be independent of the order of heat and radiation treatments.

These and further results are consistent with the view that wave lengths shorter than 2200 Å and heat act directly on the respiratory and stain resisting mechanisms. In preventing the normal increase in respiration λ2650 inactivates nucleoproteins responsible for the production of one or more links in the enzyme chain. That irradiation of cells with λ2650 sensitizes the cells to killing by heat suggests that some nucleoproteins may be injured (bonds broken) by irradiation, without destroying patterns for the formation of additional cellular constituents. However, patterns so weakened are readily destroyed by heat.

I. INTRODUCTION

In previous work from this laboratory on the effects of radiation on microorganisms, particular attention has been given to problems of survival and the multiplication of surviving cells in nutrient solution. Although the results were most interesting, it seemed difficult to make a uniquely satisfactory explanation of the effects produced.

In continuation of these investigations, the present work is a study of the effects of irradiation on certain physiological functions of the cell, respiration and stainability, in the hope that some explanation could be advanced, or so that at least certain possible explanations could be eliminated. The effects of waves

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 5, JULY, 1941

of various lengths were studied and it was found that they were selective in their action on the cellular constituents. In addition, the effects of lethal temperatures, alone and combined in various ways with irradiation, were studied and certain sensitizations to heat of the cellular constituents were found to be produced by irradiation. The techniques employed also made it possible to determine roughly the time required for the injury of one physiological function or cellular constituent to be reflected in injury to another. Direct evidence for the interdependence of cellular constituents and functions was therefore obtained. An attempt is made to explain these results and those of previous investigators in terms of modern chemical theories of biological activity.

This investigation was supported by a grant from the Committee on Radiation, Division of Biology and Agriculture, National Research Council. The authors also acknowledge assistance, by way of apparatus furnished, from the Graduate School, University of Wisconsin, and the Wisconsin Alumni Research Foundation.

II. General Considerations

A. The Survival of a Population of Cells Exposed to a Single Lethal Agent.—If a population of a single strain of microorganisms is exposed to a lethal agent the percentage of killed organisms increases with the time of exposure. The shape of the survival-time curve doubtless depends first of all on the nature of the chemical reaction or reactions brought about by the particular agent under consideration. Until proof to the contrary is provided it seems most desirable to assume that the same considerations of probability which determine the rates of chemical reactions determine the rates of all reactions in biological systems (Delbrück, 1940). However, biological systems differ greatly in complexity from ordinary chemical systems. The unit of biological activity, the cell, is tremendous in size compared to the ordinary unit of chemical activity, the molecule. The cell may be regarded as a great number of chemical systems all more or less closely linked together and coöperating in physiological activity, thus maintaining a balance or "dynamic equilibrium" (Lillie, 1932) between the cell and its surroundings. If one of the chemical systems is interfered with (by the action of a lethal agent) the balance between many or all the chemical systems may be disturbed; if sufficiently disturbed, repair will not set in, the changes become irreversible, and eventually "death" is said to be the result. The sharpness and rapidity with which certain biological reactions occur is quite analogous to that of certain so-called coöperative phenomena familiar in physics and chemistry. Indeed, as a system in which there is a large and steady interaction between a whole set of chemical systems (molecules) *certain reactions of a biological system, such as cell division, for example, might well be regarded as of this type.* In considering the cell as a chemical system it is therefore necessary to keep these features of its activity in mind.

The second factor which will determine the shape of the survival-time curve is the fact that the individual organisms in a population vary not only in appearance under the microscope —size, shape, position and size of vacuoles, etc.—but also as to the chemical composition, depending for any given cell on its age, the age of the culture, the nature of the medium, and a great number of other factors. Since the chemical compositions of the different cells may be assumed to be different, a population might be expected to vary in resistance to the lethal agent. It seems very likely that statistical results on lethal effects obtained with a population are therefore due to the superposition of the probabilities involved in the kinetics of the chemical reactions on the probabilities involved in their biological variation.

It would be extremely difficult if not impossible to separate the two probability factors which affect the results, for it should always be possible to account for a given survival-time curve by either factor alone—either by a suitable chemical mechanism or by postulating a suitable distribution of the cells with respect to sensitivity. Extraneous agents are known to affect the shapes of the curves, but they may be looked upon as affecting either the mechanism of the reaction or the distribution of the cells as to sensitivity.

On the basis either of chemical kinetics or of biological variability we might expect the cells which survive exposure to a lethal agent to behave in a different fashion from the controls. On the one hand, if the lethal action is due to a chain mechanism survivals would contain the products of the (unfinished) chain as well as the products of reactions induced by the lethal agent, but which do not normally lead to lethal effects. On the other hand, we should certainly expect the more resistant cells which survive treatment to be different from controls in other respects than merely their resistance to the lethal agent. It seems probable that both factors, chemical kinetics and biological variability, are responsible for the behavior of surviving cells.

B. The Lethal Action of High Temperatures.—The problem of the mechanism responsible for the lethal action at high temperatures has been recently reviewed by Belehrádek (1935). The number of plausible mechanisms is very great, suggesting at once that high temperatures affect directly a large proportion of the constituents of the cell, *i.e.*, that the action of heat is nonspecific. The temperature coefficients of the lethal action at high temperatures are very high, suggesting that the effects are due to some kind of coöperative action of a large number of interacting chemical systems. The sharp breaks so often observed in the variation of various functions with temperature might well be due in effect to changes in phase of a part of the biological system, either a melting of the various lipoid constituents or an "internal melting" (Langmuir, 1932) of essential protein giant-molecules producing a change in their activity and perhaps favoring denaturation. It is probably just this kind of action which is responsible for death due to exposure to high temperatures.

The present authors (Anderson and Duggar, 1938) have recently made a study of the growth and survival of yeast which had been exposed to lethal temperatures. It was found that cells receiving moderately severe treatment died more rapidly than controls in the absence of nutrient while those receiving severe treatment appeared to multiply. Treated cells also had a longer lag phase before logarithmic growth in nutrient than the controls. The criterion of death in these experiments was the loss of the ability of the cells to divide on nutrient agar to form visible colonies.

C. The Action of Radiant Energy.—General effects of ultraviolet irradiation on yeast have been discussed by Schreiber (1933, 1934). The action of radiant energy will, of course, be specific if the particular wave length is absorbed more strongly by one constituent than by any others. Some examples of the specific action of visible light on certain strongly absorbing substances (pigments) in living matter are: photosynthesis in which the absorbing substance is chlorophyll, vision in which the absorption is effected by visual purple, and the fruiting of certain myxomycetes induced in some manner by the absorption of light by certain pigments. So-called colorless organisms contain minute amounts of pigments (cytochromes, flavins, etc.) active in the respiratory systems, but as yet effects due to the absorption of these pigments have not been identified in the normal living tissue, although under certain abnormal conditions as result from the presence of carbon monoxide such effects are produced (Warburg, 1928).

It seems to be a characteristic of the action of light on many living systems that the effectiveness of the different wave lengths of the same intensity is proportional to the number of quanta absorbed by the system. This parallelism between the effectiveness of the radiation and the absorption spectrum has been noted in the conversion of visual purple to visual yellow (Schneider, Goodeve, and Lythgoe, 1939), the fruiting of myxomycetes (Gray, personal communication), and in the denaturation of proteins by ultraviolet light (*e.g.,* Kubowitz and Haas, 1933). Such a parallelism implies that the quantum efficiency of each of the processes is a constant over a wide range of wave lengths. Jordan (1938) has explained this on the theoretical basis (Möglich and Schön, 1938) that electrons are free to move about in certain energy bands in these absorbing giant molecules just as they are in solids. The radiant energy is supposed to cause an electron initially in the ground state to jump to a higher energy band and then immediately fall to the lower edge of the upper band, transforming the energy derived from this process either into light (fluorescence) or into vibration (heat). The final activated state which initiates the reaction in question is thus independent of the energy of the quantum absorbed and so the quantum efficiency of the reaction is constant over a range of wave lengths.

Wave lengths between 3100 Å and 2800 Å are known to affect the formation of vitamin D from ergosterol and to this extent may be beneficial to certain higher organisms. As yet, however, it has apparently not been shown that effects of these wave lengths on microörganisms are due to the formation of vitamin D or that the effects of shorter wave lengths are due to its destruction (Teindl-Czech, 1937).

While most of the organic cellular constituents (proteins, fatty acids, sterols, etc.) have some absorption throughout the ultraviolet, the principal absorbing substances in the region λ2800–2300 are apparently the nucleic acids which have a maximum absorption at λ2650 (Caspersson, 1936). This fact is attested by a number of observations. The relative sensitivity of cells to constant intensities of monochromatic ultraviolet radiation is a maximum at λ2650 both with regard to lethal action (Gates, 1930; Hollaender and Duggar, 1936) and with regard to the production of mutants in *Sphaerocarpus donnellii* (Knapp, Reuss, Risse, and Schreiber, 1939), both of which functions are seated in the nucleus where normally the nucleoproteins are, of course, concentrated. The fact that this wave length region is specific for the nucleoproteins opens a promising field of study, for it should be possible to investigate the function of the nucleoproteins by a study of the effects these wave lengths have upon microörganisms.

Hollaender and Duggar (1938) have studied the effects of λ2650 on *Escherichia coli* and *Serratia marcescens* particularly with respect to the subsequent growth of the organisms in nutrient. They have also studied (not yet published) the effects of this wave length on the growth and survival of *Saccharomyces cerevisiae* following treatment, and this study has been continued by the present authors (Anderson and Duggar, 1938) and extended to include the effects of heat. It is found that the survivors divide more slowly than controls in the presence of nutrient and in certain cases they die more rapidly than controls in the absence of nutrient. Also, it has been established that as a temporary effect, those cells which survive large amounts of radiation *seem* to multiply for a short time even in the absence of nutrient. In nutrient a much larger fraction of the surviving cells *appear* to divide in the first few hours of incubation than do the controls. This preliminary increase is followed by a prolongation of the lag phase of growth over that of the controls.

For wave lengths shorter than 2300 Å the absorption of the proteins as well as of other substances increases greatly, so that the effects produced by these wave lengths are probably nonspecific. Since the absorption of the protoplasm is high, however, there is some chance that the surface structures of large cells absorb most of the energy and are greatly affected, at the same time shielding the material at the center of the cell from the action of the light. These short wave lengths might thus be used to alter the surface properties of the cell. Apparently no specific attempts to make such a study have been made thus far.

Ionizing radiations such as X-rays, β-rays, etc. also attack uniformly all the molecules in living matter because in the case of X-rays the primary interaction is with the inner (non-binding) electrons of the atoms, and second, the energies involved are great enough to break any chemical bond. It is precisely because of this uniform action on cellular material that the effects of ionizing radiations are expected to be roughly proportional to the volume of the element in the cell which must be hit to produce death or to inhibit a function of a cell (Lacassagne, 1934). At least one precaution must be observed in interpreting results of experiments along these lines, however; highly active molecules and fragments are produced by the action of ionizing radiations which in a secondary process might destroy certain constituents of the cell (Failla, 1937). Effects produced by such a process would not necessarily be proportional to the volume of the "sensitive spot," but might well depend on entirely different factors, such as the composition of the medium surrounding the cell as related to the composition of the "sensitive spot." The observations of Duryee (1939) might be noted in this connection. He reported that X-irradiation of frog egg chromosomes removed from the surrounding cytoplasm produced no effect while the irradiation of the chromosomes in cytoplasm resulted in injury to the chromosomes.

That some kind of recovery follows irradiation with X-rays is shown by the work of Henshaw (1938) who noted that the delay in the first cleavage after immediate insemination of irradiated Arbacia eggs does not occur. However, the injury persists in later cleavages (Henshaw, 1939).

D. The Action of Two Lethal Agents on a Population of Cells.—When a population of cells is treated with two lethal agents the shape of the survival-time curves will depend on the same considerations mentioned in II *A*. In addition, however, the results will depend on the *relations* between first, the kinetics of action of the two agents and second, the separate sensitivity distributions of the cells to the two agents. These relations may be of certain types—and given certain types of relationships it is possible to predict the kind of result to be obtained.

Let us consider as an example the effects of combining heat and radiation treatments on a certain physiological property of a suspension of cells. Let us imagine that one sample *H* from this suspension is treated with heat and another sample *R* with radiation. Then let us take part *HR* of the heat treated sample and give it a treatment with the same amount of radiation as the *R* sample received and then give part *RH* of the *R* sample the heat treatment which *H* received. We will then have five samples for study; the control *C*; the two samples *H* and *R*, given heat and radiation treatments, respectively; and the two samples HR and RH, given combinations of identical heat and radiation treatments, but combined in reverse order. We shall be interested in measuring the extent to which a physiological activity of the suspension, the rate of respiration, say, is diminished by each of the treatments. Let the fraction by which the activity of the function is reduced be *h*, *r*, *hr*, and *rh* for each of the suspensions *H*, *R*, *HR*, and *RH*, respectively (the fraction for the control *C*, being of course 1). The relations between the various fractional reductions in activity on various hypotheses may then be set forth in outline form.

1. Heat and radiation act on *different and independent* functional units[1] in the cells; $rh = hr$.
    - *a.* If cells are identical, or if there is no correlation between sensitivities of the cells to heat and radiation; $rh = hr = r \times h$ (where $r \times h$ signifies ordinary multiplication of the numbers r and h).
    - *b.* If cells which are more sensitive to heat are also more sensitive to radiation, $rh = hr > r \times h$.
    - *c.* If cells which are more sensitive to heat are less sensitive to radiation, $rh = hr < r \times h$.
2. If heat and radiation act on different but related functional units which are involved in the function under study we

---

[1] For want of a better term we use the expression "functional unit" for that molecule or complex of molecules in the cell which is directly responsible for the physiological or biological process under study.

should expect the closeness of this relation to be reflected in the time required for the effect of one treatment on the other to become apparent. Thus, the functions expressing the relations between r, h, rh, and hr would involve the time after the treatment t; thus $rh = f_1\ (r, h, t)$, and $hr = f_2\ (r, h, t)$ where f_1 and f_2 are, in general, different. The closeness of the relation between the functional units affected by the two agents would probably be reflected in the extent to which f_1 and f_2 depend on the time t.

- *a.* A special case would be that in which one of the lethal agents, radiation, say, produced substances in the cell which made the contents then more susceptible to heat treatment; thus $rh<hr$.
- *b.* Another would be that in which one of the agents radiation, say, produced a substance in the cell which made it then less susceptible to the action of the other; thus $rh>hr$.

3. Radiation and heat act on the same functional units.
   - *a.* The mechanisms of action of the two treatments are in effect the same; thus $rh = hr$.
     - (1) The cells are identical.
       - (*a*) If one molecule only needs to be inactivated to inactivate the function under study (a monomolecular reaction), $rh = hr = r \times h = e^{-(at_r + bt_h)}$ where t_r and t_h are times of radiation and heat treatment respectively and a and b are constants.
       - (*b*) If more than one molecule must be inactivated to inactivate the function the value of $rh = hr$ may be calculated from the relations discussed by Mme. P. Curie (1929). In general $rh = hr<r \times h$.
     - (2) If the cells are not identical the relation between rh and $r \times h$ will depend on the nature and extent of variation between the cells. The inequalities in 1.*b*. or 1.*c*. will then be superimposed on the equations in 3.*a* (1).

b. If radiation and heat act on the same functional units in different ways.

  (1) If each agent inactivates directly the mechanism responsible for the function, $rh = hr$.

  (2) If one agent, radiation, say, reacts with, but does not inactivate the mechanism in such a way that either

    (*a*) the mechanism is then made more susceptible to the other agent, thus $rh < hr = h$; or

    (b) the mechanism is then made less susceptible to the other agent, thus $h = hr < rh$.

  (3) If each agent makes the function more susceptible to the other agent, but in different degrees, radiation more than heat, say, $rh < hr < h \times r$.

If it were possible to formulate a general theory according to the methods outlined in 2., 1. and 3. would be special cases in which, among other things, $t = \infty$ and $t = o$, respectively. It is, of course, possible to formulate many more complicated mechanisms than those above, but these seem to be the simplest and we will arbitrarily limit the discussion to them for the time being.

Bovie and Daland (1923) have observed that *Paramecium* is sensitized to heat by irradiation with fluorite rays ($\lambda < 2000$ Å). Their results, together with those obtained in the present work, will be discussed later from the above point of view.

E. Scope and Purpose of the Present Work.—As has already been pointed out, the effects of irradiation and heat on the survival and growth of yeast have been studied. However, it was felt that the determination of such effects by the normal plating out technique allows considerable time for the physiological functions which are upset by the lethal agent to interact with and destroy the dynamic equilibrium existing between the various functions. Colony formation could thus appear as a net effect and the interpretation of results in terms of specific physiological functions would be difficult. For this and other reasons it was thought that a deeper insight into the mode of action of heat and radiation might be gained by observing the effects of treatment on certain physiological activities of the cell which could be conveniently observed with precision during or immediately following treatment. In this way, too, knowledge as to the relationships existing between the various physiological functions of the cell might be obtained.

In this paper are presented the results of such an investigation. Methods are developed to determine the effects of treatment on the following properties of yeast: (*a*) the resistance of the cell to staining with methylene blue, (*b*) the rate of respiration of cells in a nutrient solution, and (*c*) the rate of growth in the number of cells as determined by increases in the rate of respiration and cell counts. In this manner the relative sensitivities of each of these functions to heat and various wave lengths of ultraviolet light have been established and indications have been obtained as to how they vary with time after exposure. It also appeared highly desirable to determine which of the functions are affected directly by both treatments as well as to determine the possible relations existing between functions apparently affected by one treatment but not by the other. A study of the effects of combined treatments with heat and $\lambda 2650$ was therefore made. This mode of attack also has led to a formulation of the nature of the chemical reactions leading to death which are brought about by heat and different wave lengths of ultraviolet light.

III. EXPERIMENTAL METHODS

A. Yeast Culture.—In certain preliminary experiments ordinary commercial Fleischmann's yeast cake was used. In all cases in which plates were poured and in all final experiments the experimental organism was obtained from a single cell isolation of a Burgundy wine strain of *Saccaromyces cerevisiae.* For use the yeast was cultured for 42 hours on slants of potato dextrose agar made up of the following constituents: broth from 200 gm. of potatoes boiled for 60 min. in 500 cc. of water, 15 gm. of agar dissolved in 400 cc. of water and 0.2 gm. $MgSO_4$, 0.2 gm. $CaCO_3$, and 20 gm. of dextrose dissolved in 100 cc. of water.

B. Preparation of Suspensions.—Two solutions for the suspension of the yeast were used in this work. A "physiological" salt solution had the following composition: 0.3 per cent sodium chloride, 0.02 per cent potassium chloride, and 0.02 per cent calcium chloride. The "nutrient" solution contained 0.5 gm. KH_2PO_4, 0.2 gm. $MgSO_4$, 105 gm. dextrose, 1.0 gm. asparagin. 3.0 gm. NaCl, 20.2 gm. KCl and 0.2 gm. $CaCl_2$ in 2 liters of distilled water. This nutrient solution had the advantage of being transparent to the ultraviolet to 2300 Å as well as providing sufficient electrolyte for the proper behavior of the dropping mercury electrode which was used for the determination of oxygen in the cell suspension.

Ordinarily, yeast was washed from the slants with the solution in which it was to be treated. The suspension so obtained was then washed twice by centrifuging and resuspending in fresh solution. In certain cases in which the rate of respiration of a suspension was to be measured it was desirable to have yeast which would remain in suspension for rather long periods of time. In these cases the yeast was usually centrifuged again for 30 sec. and the supernatant suspension of cells was used for study. This suspension was then filtered through filter paper to break up clumps of cells, and what was more important, to add certain substances from the paper which somehow insure the proper behavior of the dropping mercury electrode used to determine the rate of respiration (Petering and Daniels, 1938).

C. Heat Treatment.—For the exposure of the suspensions to heat, a few cc. of the yeast suspension were placed in a standard centrifuge tube of 15 cc. capacity, stoppered to prevent evaporation. The tube was placed in a water thermostat, the temperature of which was maintained at 50 ± 0.02° C. as determined by a thermometer calibrated by the National Bureau of Standards. This temperature had been found sufficiently effective to inactivate the physiological functions of the yeast rapidly and yet slow enough in its action to permit the accurate determination of the time of treatment. Rapid stirring of the thermostat and the small volume of the suspension compared with the surface exposed to the water tended to make the suspension come to constant temperature rapidly. Likewise, when the tubes were removed from the thermostat they came to room temperature quickly. The time of exposure of the tubes to the thermostat water was found to be a sufficiently accurate measure of the severity of the treatment.

D. Radiation Treatment.—In this work the ultraviolet light was obtained from a Type H–3 General Electric mercury vapor lamp. The outer glass bulb surrounding the quartz tube inside of which the high pressure discharge takes place was replaced by a metal shield with a small vertical slit to permit the exit of the shorter wave lengths. This light was focused on the slit of a quartz monochromator which served to isolate a given wave length when so desired. In other more qualitative experiments liquid filters were used as described later.

In certain cases the yeast suspension was irradiated in a small quartz optical cell placed in the beam of light from the monochromator which was large enough to illuminate the surface of the cell uniformly. Dilute suspensions were used so that only a small percentage of the light was absorbed by the suspension and each yeast cell was therefore exposed to approximately the same amount of illumination. In other cases it was possible to irradiate the yeast in the quartz cell of the dropping mercury electrode while the rate of respiration of the suspension was being measured. This quartz cell was also small enough to fit into the beam of the monochromator.

E. Measurement of the Rate of Respiration.—A number of classical methods for the determination of oxygen are available, but that which is most convenient and which, when properly used, gives most accurate results is the dropping mercury electrode, recently developed for this purpose in this laboratory by Petering and Daniels (1938). A number of extremely simple electrical setups are given in detail in their paper; the setup used in this work is essentially that in which a potentiometer is used as a source of e.m.f. A cylindrical quartz cell about 2 cm. in diameter and 1 cm. thick was used for the oxygen determinations in the present paper. With the dropping mercury electrode it happens that the difference between the currents (measured by the galvanometer deflections) which flow through the cell with 0.1 V. and 1.0 V. applied e.m.f. is very nearly proportional to the oxygen concentration of the solution. The rate of change in the oxygen concentration is then a measure of the rate of respiration of the cells in the suspension.

F. Staining by Methylene Blue.—The procedure adopted by Fulmer and Buchanan (1923) in their study of effects of toxic agents was followed in this work. For this purpose 1 cc. of the suspension of cells to be tested was added to 1 cc. of 0.01 per cent of methylene blue and the percentage of stained cells determined by microscopic observation of the cells in a Levy counting chamber. The counts were made after two hours of exposure of the cells to the solution, and usually again after an exposure of a day or more.

G. Determination of Concentration of Colony-forming Cells. —The concentration of colony-forming cells in a suspension was determined by the customary dilution and plating out technique using the agar already described for pouring plates. When this was done care was taken, of course, to maintain sterile conditions to prevent contamination of the plates.

IV. Experimental Results

A. The Regions which are Effective in the Ultraviolet.—A number of orienting experiments were made with the yeast suspended in physiological salt solution. When this suspension in a fused quartz test tube was exposed to the full arc at a distance of 15 cm. few cells took the methylene blue stain after 1 min. exposure (enough to prevent 99 per cent of the cells from forming colonies) and the rate of respiration was considerably reduced. However, after 5 min. exposure 99 per cent took the stain. To see whether substances formed in the solution, such as ozone, might be responsible for the results, physiological salt solution was irradiated with the full arc for 30 min. and yeast cells then added to it. They did not take the stain. A suspension of cells was then placed in a test tube partially covered with black paper so as to shield the suspension from the light but leave the air above the suspension exposed. After 25 min. irradiation at 15 cm. distance the contents of the tube were shaken and the staining test made. Again no cells took the stain. Evidently in this case the action of the arc is not due to the production of stable toxic substances in the physiological salt solution nor to ozone in the air above it (Bovie and Daland, 1923). A suspension of control cells was added to a suspension which had been irradiated for 25 min. with the full arc. The control cells did not take the stain. It seems most reasonable to assume as a result of these experiments that the light affects the contents of the cells directly and that the products formed in exposed cells do not affect normal cells, at least to the extent of making them stain with methylene blue.

A suspension of cells in physiological salt solution was irradiated with the full arc filtered through ethyl alcohol to absorb wave lengths shorter than 2200 Å. The cells did not take the stain even after a 30 minute exposure, nor was the rate of respiration reduced appreciably. This exposure was sufficient to prevent 99 per cent or more of the cells from forming colonies on agar. Evidently the region which is most effective in producing changes resulting in staining and also in reducing respiration is absorbed by alcohol, *i.e.* is shorter than 2200 Å. The activity is probably due principally to the intense pair of mercury lines at 1942 Å and 1849 Å. Later it will be shown that prolonged exposure to λ2650 does affect staining and respiration to an extent which is slight, but negligible compared with the inactivation of the colony-forming ability of the cells.

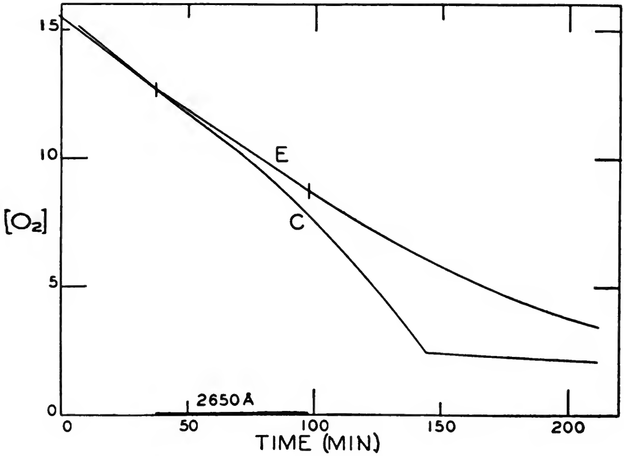


FIG. 1. The effect of irradiation of yeast with λ2650 on its oxygen consumption (Curve *E*) as compared with that of the control (Curve *C*).

B. Effects of Radiation on the Rate of Respiration.—In this section are given first the results of experiments in which the yeast suspended in nutrient was irradiated in the quartz cell of the dropping mercury electrode. The advantage of this procedure is that the rate of respiration of the same cell suspension may be studied before, during, and after the irradiation. Thus the rapidity with which the effects of radiation are produced can be easily determined. The principal disadvantage of the method is that the time during which the effects of radiation can be studied is limited by the eventual depletion of the oxygen supply and the gradual settling of the cells in the suspension.

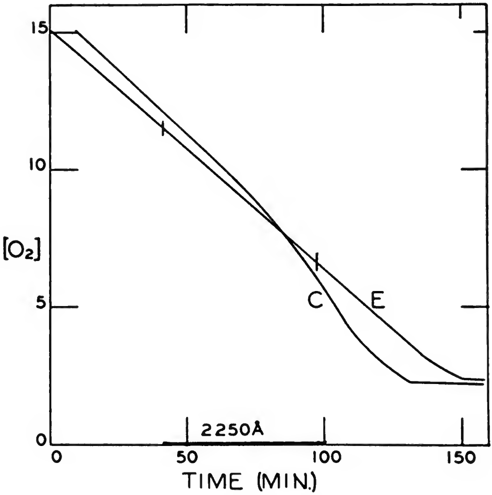


Fig. 2. The effect of irradiation of yeast with λ2250 on its oxygen consumption (Curve *E*) as compared with that of the control (Curve *C*).

A number of experiments were performed in which the rate of respiration of an exposed suspension was compared with that of a control suspension in a second dropping mercury electrode cell. In Fig. 1 the results of such an experiment are given. Here the oxygen concentrations are plotted in arbitrary units (galvanometer deflections) as ordinates against the times as abscissæ. The rate of respiration (proportional to the slope of the curve) of the suspension exposed to λ2650 (Curve *E*) for 60 minutes decreases slowly after the exposure. That of the control (Curve *C*) is seen to increase steadily after 50 minutes until after 150 minutes the oxygen supply is exhausted and the deflection of the galvanometer becomes constant. The same result was obtained with a 21 minute exposure to λ2650. A 56 minute exposure to a low intensity of λ2250 (Fig. 2) also prevents the normal increase in respiration. The results of an experiment in which yeast in nutrient solution was irradiated first with λ2650 and then with the full arc are given in Fig. 3. It is

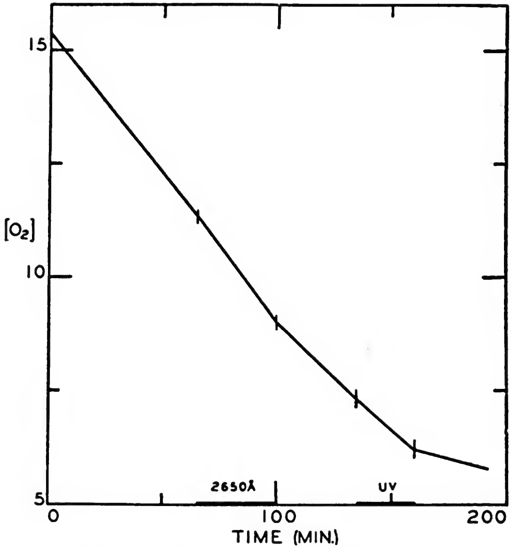


Fig. 3. The effect of irradiation of yeast with λ2650 and the full mercury arc (U.V.) on its consumption of oxygen.

to be seen from the figure that the rate of consumption of oxygen remains practically the same during the 35 minutes of irradiation with λ2650 as before, but falls off appreciably after the irradiation is stopped. The same is true during the period of irradiation with the full arc, but after this irradiation the rate of consumption of oxygen becomes negligible; after 1000 minutes an appreciable amount of oxygen was still found in the cell.

This result, which was also obtained in a number of additional experiments, was surprising, because if radiation inhibits the respiratory mechanism in yeast one would expect the rate of consumption of oxygen to fall off during the period of irradiation. It appeared that this extra consumption during irradiation might have been due to an oxidation reaction analogous

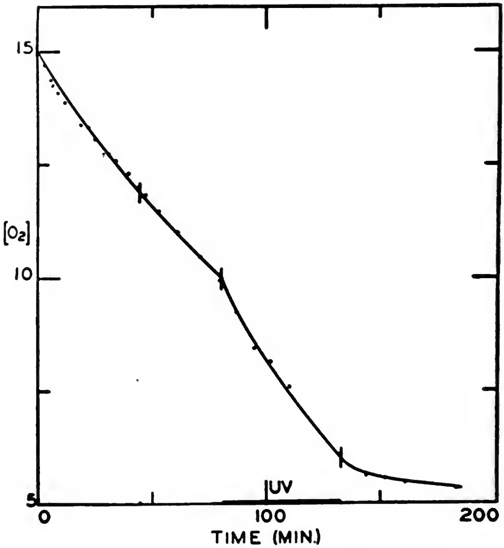


FIG. 4. The effect of irradiation of yeast with a high intensity of the full arc on the consumption of oxygen.

to that which is known to follow the photo-reaction in the denaturation of proteins by ultraviolet light (Mitchell and Rideal, 1938). According to this picture the rate of oxygen consumption during irradiation would be determined by two factors, the rate of normal respiration (decreasing during the exposure) and the rate of the oxidation reaction following the photo-reaction (increasing during exposure and continuing after the exposure). That the rate of oxygen consumption should remain constant during exposure, as it does in the experiment of Fig. 2 and in other experiments, would appear to be purely accidental.

To test this possibility further a number of additional experiments were carried out. In one experiment illustrated in Fig. 4, the yeast was irradiated with a very high intensity of radiation from the full arc. It is seen that the rate of oxygen uptake in this experiment was *greater* during the period of irradiation than before. This can be interpreted as indicating that the oxidation reaction has been speeded up by the high intensity of light until it *exceeds* the normal rate of respiration. Apparently this photo-reaction continued for some minutes after the irradiation was terminated, becoming practically zero after a period of ten minutes.

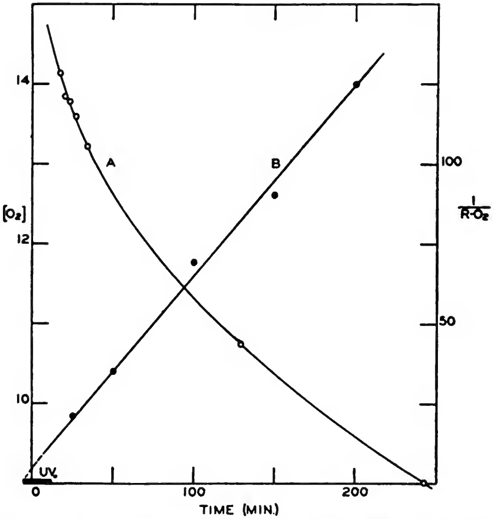


Fig. 5. Curve *A*. The oxygen consumption of a yeast which had been irradiated with the full arc at room temperature. Curve *B*. The rate of oxygen consumption times the oxygen concentration is inversely proportional to the time.

It appeared feasible to follow the course of this oxidation reaction itself by irradiating the yeast suspension with very high intensity for a relatively short time and then following the oxygen uptake as a function of the time. A few preliminary experiments of this nature were therefore carried out.

Yeast from Fleischmann's yeast cake was suspended in physiological salt solution and a portion was irradiated in a quartz test tube for 18 minutes with the very intense ultraviolet radiation from a capillary mercury arc of the type described by Daniels and Heidt (1932). One cc. of the remaining control suspension was added to 10 cc. of nutrient solution and the rate of respiration found to be 0.95 in arbitrary units. The irradiated suspension was diluted in the same manner and the rate of oxygen consumption similarly measured with the results shown in Fig. 5 (Curve *A*). The rate of respiration *R* of the exposed culture fell off steadily with time from 0.13 to 0.015. It

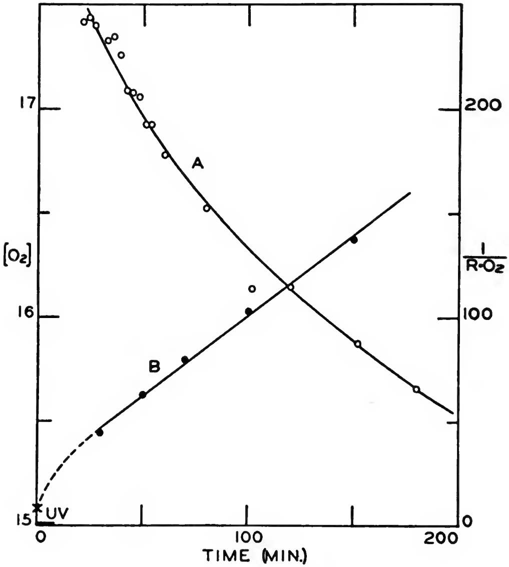


FIG. 6. The results of an experiment patterned after that of Fig. 5, but in which the yeast was kept at 4° C.

was found empirically that the reciprocal of the product of the rate of respiration and the relative oxygen concentration $1/R \cdot O_2$ is proportional to the time (Curve *B*, Fig. 5). The curve extrapolates back to the value of $1/R$ for the control at the start of the irradiation.

It was thought that temperature might affect the rate of the oxidation reaction, particularly during the irradiation. The method of the above experiment was therefore duplicated but here the yeast was kept at 4° C. during and after the irradiation until it was diluted in the nutrient solution and the rate of oxygen consumption measured at room temperature. The results shown in Fig. 6 are roughly the same as in the previous experiment, showing that the rate of the reaction which is responsible for the reduction in *R* is relatively independent of temperature in this range.

That the wave lengths in the spectrum of the capillary arc which are responsible for the rapid decrease of *R* are shorter than 2200 Å is shown by the following experiment. The yeast suspended in physiological salt solution in a small quartz test tube was irradiated for 31 minutes with the full arc filtered through ethyl alcohol contained in a large quartz test tube placed outside that containing the yeast. The rate of respiration of the control suspension and the exposed suspension both diluted in nutrient as above were then measured and found to be equal within 5 per cent. Next day, however, the respiration of the exposed yeast (which had been kept in physiological salt solution at 4° C.) had dropped 90 per cent while that of the control suspension remained constant. It is to be concluded from this experiment that alcohol absorbs most of the radiation which is immediately effective in reducing the rate of respiration. The most effective wave lengths are therefore less than 2250 Å, the absorption limit of alcohol. The effective absorption is probably by molecules of the respiratory system of yeast, but it is possible that products are produced by irradiation with these short wave lengths which may "poison" the respiratory enzyme system.

C. The Effects of Combining Heat with Irradiation at λ2650. —As mentioned in II.-*D*. special interest would attach to an investigation of the effects of combinations of heat and ultraviolet treatments, most particularly ultraviolet of wave lengths 2650 Å which would affect principally the nucleic acid. Consequently a set of experiments of the following general plan was carried out. One portion *R* of a suspension of standard yeast in physiological salt solution was exposed for a certain length of time to λ2650. Another portion *H* was exposed to a temperature of 50° C for another length of time. A portion *RH* of the irradiated suspension was then exposed to 50° C. and a portion *HR* of the heat treated suspension was irradiated with λ2650. The untreated control suspension *C*, the two suspensions *R* and *H* given separate treatments, and the two suspensions *RH* and *HR* given equivalent treatments combined in reverse order thus gave five samples available for study. Dilution plates were poured and determinations of the per cent of stained cells were made on each sample. One cc. of each sample was added to 10 cc. of the nutrient solution and the rates of respiration were measured. It was soon found, however, that the sensitivities of the three physiological processes to the various treatments were so different that it was impractical to attempt to make determinations of all three of these processes in each case.

The results of nine experiments of varying severity of treatment are given in Tables I and II. In the first column of the table is given the number of the experiment. The samples are listed in the second column, the number following the letter *R* or *H* indicating the time in minutes of radiation or heat treatment given the sample. Thus the notation R5H10 indicates a sample which had been exposed to λ2650 for 5 minutes and then to 50° C. for 10 minutes. In the third column of Table I is given the concentration of colony-forming organisms in each sample for moderate treatments and in Table II, the relative rates of respiration for more severe treatments and the values calculated from Eq. (1). The survivals as determined from colony counts are given in the fourth column together with values calculated with Eq. (1), while the fifth column contains the percentage of cells which resisted staining by methylene blue after 1 hour of exposure to the dye. The sixth column contains the percentage of cells which resisted staining after longer periods of exposure to the dye.

The ability of the cells to form colonies is seen to be the most sensitive of the three physiological processes both to heat and irradiation with λ2650. Respiration is quite sensitive to heat,

TABLE I

BEHAVIOR OF YEAST EXPOSED TO λ2650, TO 50° C. AND TO COMBINED TREATMENTS: COLONY FORMATION AND STAINING

No. of Experiment	Sample	Concen. of Colony-forming Cells	Survival Per Cent	Per Cent of Stain Resisting Cells	
				After 1 hr.	Later
I.					(28 hr.)
	C	78×10^6	100	99.1	96.4
	R5	52×10^6	92	98.9	97.9
	H10	72×10^6	67	96.4	89.0
	R5H10	133×10^5	17	92.1	79.2
	H10R5	193×10^5	25	97.0	88.9
Calc.	H10R5		62		
II.					(46 hr.)
	C	276×10^5	100	99.4	91.6
	R7	738×10^4	26.7	99.1	94.5
	H15	278×10^4	10.1	98.9	15.8
	R7H15	137×10^2	0.050	95.4	6.9
	H15R7	322×10^2	0.117	97.9	41.5
Calc.	H15R7		2.6		
III.					(44 hr.)
	C	163×10^5	100	99.1	67.2
	R7	228×10^4	14.0	99.5	54.8
	H15	748×10^4	45.8	97.9	45.7
	R7H15	145×10^3	0.089	97.7	63.0
	H15R7	69×10^3	0.42	98.0	41.5
Calc.	H15R7		6.4		
IV.	C	182×10^5	100.0		
	R10	166×10^3	0.91		
	H25	101×10^3	0.56		
	R10H25	4×10^2	0.002		
	H25R10	15×10^2	0.008		
V.					(19 hr.)
	C			98.1	95.8
	R15			99.0	92.3
	H30			97.6	6.3
	R15H30			92.5	23.9
	H30R15			97.1	2.3

but is only slightly affected by 45 minutes of irradiation with λ2650 (Exp. IX). Likewise the resistance of the cells to staining is reduced to a greater extent by heat than by λ2650; only in Exp. VII in which the suspension of cells was irradiated for 60 minutes is there a significant difference between *R* and *C*

TABLE II

Behavior of Yeast Exposed to λ2650, to 50° C. and to Combined Treatments
Rate of Respiration and Staining

No. of Experiment	Sample	Relative Rate of Respiration		Per Cent of Stain Resisting Cells	
		After 1 hr.	Later	After 1 hr.	Later
VI.					(45 hr.)
	C	0.199		97.5	94.1
	R20	0.204		98.2	53.0
	H20	0.072		93.7	13.6
	R20H20	0.064		92.8	18.4
	H20R20	0.059		96.5	13.4
Calc.	H20R20	0.074			
VII.					(21 hr.)
	C			98.7	87.5
	R30			98.7	88.4
	H60			95.1	7.8
	R30H60			85.1	52.3
	H60R30			97.9	26.7
VIII.					(45 hr.)
	C			99.8	82.6
	R60			91.1	68.0
	H120			90.5	2.1
	R60H120			20.3	0.6
	H120R60			65.5	0.6
IX.			(7 hr.)		
	C	0.400	0.400		
	R45	0.350	0.330		
	H20	0.248	0.285		
	R45H20	0.228	0.234		
	H20R45	0.231	0.235		
Calc.	H20R45	0.218	0.236		

after one hour of exposure to the dye. The observations after longer exposure to the dye are less consistent, but here too, heat treatment produced more staining than radiation in all experiments.

It is significant that in only one case (Exp. III, Sample H15) did the percentage of cells which took the methylene blue stain after 2 days of exposure correspond to the percentage which failed to produce colonies.[1] In all other treated samples the percentage of cells which took the stain after two days exposure was always considerably less than that which failed to produce colonies, in spite of the fact that the dye itself is toxic, witnessed by the large percentage of control cells which eventually took the stain after two days of exposure to the dye. These facts together with the great differences in sensitivity of the two processes to lethal agents argue strongly for the point that the two processes depend on quite separate mechanisms for their expression. Also, the fact that a cell resists staining does not mean that it is "living" in the sense that it would be able to multiply to produce a colony on nutrient agar. As far as is known, however, a cell which does stain would be incapable of producing a colony if placed in a favorable medium.

---

[1] There is a possibility that the effects with methylene blue might be complicated by the reduction of the dye to a colorless form. Further experiments will be conducted using a non-reducible dye.

Probably the most striking effect to be noted from Table I is the sensitization to heat treatment caused by irradiation with λ2650. In every case the concentration of colony-forming organisms in the *RH* sample is much lower than that in the *HR* sample. Similarly, a much smaller percentage of the cells in the *RH* sample resist methylene blue stain after an hour of exposure to the dye than in the *HR* sample. Irradiation with λ2650 thus sensitizes to heat treatment both the colony-forming ability and resistance to staining.

The effects of heat and radiation are not additive, however, even when radiation follows heat treatment. The additive relation

$$hr = h \times r \tag{1}$$

of II.-*D.* was used to calculate the values for the survivals in Table I. It is seen that the observed values for the *HR* sample are considerably lower than the calculated values (except in Exp. IX where the numbers of colonies counted were too small to be significant).

Neither are the effects additive with regard to staining relations. It is interesting to note that in a considerable number of experiments (I, III, V, VI, VII) radiation appears to reduce the amount of staining either of the control or of the heat treated sample. The results obtained after long periods of exposure to the dye appear to be rather erratic. Since a long time is allowed in these experiments between exposure and the making of observations, complicated interactions may have occurred giving rise to the high apparent variation.

In contrast with the above, the rate of respiration of the *HR* and *RH* samples are equal within experimental error; λ2650 does not sensitize the respiratory system to heat. Also, from the agreement between the calculated and observed values it is apparent that the effects of λ2650 and of heat are additive in their effect upon respiration, at least within the range of values studied. The simple case in which both of these relations should exist is that [II.-*D*. 3.*a*.(1)(*a*)] in which heat and radiation act on the same functional unit in the cell and each independently can inactivate the mechanism.

V. Conclusions

Wave lengths shorter than 2200 Å only are effective in producing immediate staining of yeast with methylene blue and the effects appear to result from the direct action of the light on the cellular constituents.

Previous irradiation of yeast with λ2650 prevents the immediate gradual *increase* in respiration which normally occurs when the yeast is placed in nutrient. The oxygen uptake of the cell during irradiation with high intensities of ultraviolet light depends on (*a*) the rate at which normal aerobic respiration is reduced, and (*b*) the rate of photoöxidation of the cellular constituents which continues after the irradiation has ceased; all three processes depend, of course, on the light intensity.

Irradiation of yeast with λ2650 sensitizes only slightly the cells to immediate staining with methylene blue. The killing (removal of colony-forming capacity) of cells by heat is sensitized to a marked degree by previous irradiation with λ2650. On the other hand, heat treatment and irradiation with λ2650 are additive with respect to decreasing the rate of respiration. This result implies that radiation and heat act independently on the rate determining links in the enzyme chain.

It seems desirable to attempt to interpret briefly the results which involve growth in terms of modern theories of biological activity (Jordan, 1939; Pauling and Delbrück, 1940) in which certain molecules in the living cell are looked upon as patterns active in attaching groups to themselves, reproducing in this way the specific molecules characteristic of the species and necessary for its propagation.

The prevention of the normal increase in respiration by irradiation with λ2650 suggests that one or more of the patterns necessary for the synthesis of respiratory enzymes may be broken by light of this wave length. Since the principal cellular constituents which absorb this wave length are the nucleo-proteins, the implication is that the nucleoproteins are in some way responsible for the production of at least one of the links in the enzyme chain.

Radiation alone and heat alone are, of course, capable of breaking patterns. That irradiation of cells with λ2650 sensitizes the cells to killing by heat suggests that the nucleo-proteins may be injured (bonds broken) by irradiation, but that they are still capable of functioning as patterns to some extent for the formation of additional cellular constituents, unless the weakened molecules are further denatured by heat (Clark, 1925). From the results on oxygen uptake during and following intense irradiation, it appears that one step involved in the photo-denaturation is an oxidation.

Other explanations of these effects are, of course, possible, but we believe that the essential explanation of the effects of light and of heat must involve either the direct inactivation of functional units including the specific patterns, or the production of products which, in a secondary reaction, inactivate the functional units or patterns.

VI. REFERENCES

ANDERSON, T. F., AND DUGGAR, B. M. 1938. Effects of heat on yeast. I. Rate of death of cells which survive moderate treatment. *Amer. Journ. Bot.*, **25**, 11 s.

BĚLEHRÁDEK, J. 1935. Temperature and living matter. *Protoplasma-Monographien*, **8**, 277 pp. Berlin. (See p. 48.)

BERNHART, F. W. 1939. Changes in pH of irradiated egg albumin solutions. *Jour. Amer. Chem. Soc.*, **61**, 1953–1954.

BOVIE, W. T., AND DALAND, G. A. 1923. New experiments on the sensitization of protoplasm to heat by exposure to light of short wave-length. *Amer. Jour. Physiol.*, **66**, 55–89.

CASPERSSON, T. 1936. Über den chemischen Aufbau der Strukturen des Zellkernes. *Skand. Arch. Physiol.*, **73** (suppl. 8), 1–151.

CLARK, J. H. 1925. Studies on radiated proteins. I. Coagulation of egg albumin by ultra-violet light and heat. *Amer. Jour. Physiol.*, **73**, 649–660.

CURIE, MME. P. 1929. Sur l'étude des courbes de probabilité relative à l'action des rayons X sur les bacilles. *Compt. Rend. Acad. Sci. Paris*, **188**, 202–204.

DANIELS, FARRINGTON, AND HEIDT, L. J. 1932. A simple capillary mercury vapor lamp. *Jour. Amer. Chem. Soc.*, **54**, 2381–2387.

DELBRÜCK, M. 1940. Statistical fluctuations in autocatalytic reactions. *Jour. Chem. Phys.*, **8**, 120–124.

Duryee, W. R. 1939. Does the action of X-rays on the nucleus depend upon the cytoplasm? *Biolog. Bull.*, **77**, 326.

Failla, G. 1937. A theory of the biological action of ionizing radiations. (In "Some fundamental aspects of the cancer problem.") *Occas. Publ. Amer. Assoc. Adv. Sci.*, No. **4**, pp. 204–214.

Fulmer, E. I., and Buchanan, R. E. 1923. Studies on toxicity. *Jour. Gen. Physiol.*, **6**, 77–89.

Gates, F. L. 1930. A study of the bactericidal action of ultraviolet light. III. *Jour. Gen. Physiol.*, **14**, 31–42.

Henshaw, P. S. 1938. The effect of X-rays on *Arbacia punctulata* sperm. *Biolog. Bull.*, **75**, 347.

Henshaw, P. S. 1939. Fixation of X-ray effect by fertilization in Arbacia eggs. *Biolog. Bull.*, **77**, 325–326.

Hershey, A. D. Factors limiting bacterial growth. *Jour Bacteriol.*, **38**, 563–578.

Hollaender, A., and Duggar, B. M. 1936. Irradiation of plant viruses and of microorganisms with monochromatic light. III. *Proc. Nat. Acad. Sci.*, **22**, 19–24.

Hollaender, A., and Duggar, B. M. 1938. The effects of sublethal doses of monochromatic ultraviolet radiation on the growth properties of bacteria. *Jour. Bacteriol.*, **36**, 17–37.

Jordan, P. 1938. Über die physikalische Struktur organischer Riesenmoleküle. *Naturwiss.*, **26**, 693–694.

Jordan, P. 1938. Zur Frage einer spezifischen Anziehung zwischen Genmolekülen. *Phys. Zeits.*, **39**, 711–714. (See also articles cited by Pauling and Delbrück, 1940.)

Knapp, E., Reuss, A., Risse, O., and Schreiber, H. 1939. Quantitative Analyse der mutationsauslösenden Wirkung monochromatischen Uv.-Lichtes. *Naturwiss.*, **27**, 304.

Kreuschen, K. H., and Bateman, J. B. 1935. Physikalische und biologische Untersuchungen über mitogenetische Strahlung. *Protoplasma*, **22**, 243–273.

Kubowitz, F., and Haas, E. 1933. Über das Zerstörungsspektrum der Urease. *Biochem. Zeitsch.*, **257**, 337–343.

Lacassagne, A. 1934. Le problème des quanta en radiobiologie. *Jour. Radiol. et Electrol.*, **18**, 553–560.

Langmuir, I. 1932. Vapor pressure, evaporation, condensation, and adsorption. *Jour. Amer. Chem. Soc.*, **54**, 2798–2832. (See footnote 10, p. 2805.)

Lillie, R. S. 1932. Protoplasmic action and nervous action. Univ. of Chicago. (Chapt. 4.)

Mitchell, J. S., and Rideal, E. K. 1938. Photochemical reactions in monolayers. II. The photochemistry of proteins. *Proc. Roy. Soc., London*, **167A**, 342–366.

Möglich, F., and Schön, M. 1938. Zur Frage der Energiewanderung in Kristallen und Molekülkomplexen. *Naturwiss.*, **26**, 199.

Pauling, L., and Delbrück, M. 1940. The nature of the intermolecular forces operative in biological processes. *Science*, **92**, 77–79.

Petering, H. G., and Daniels, Farrington. 1938. The determination of dissolved oxygen by means of the dropping mercury electrode with applications in biology. *Jour. Amer. Chem. Soc.*, **60**, 2796–2802.

Schneider, E. E., Goodeve, C. F., and Lythgoe, R. J. 1939. The spectral variation of the photosensitivity of visual purple. *Proc. Roy. Soc., London*, **170A**. 102–112.

Schreiber, H. 1933. Zur Theorie der "mitogenetischen Strahlung." *Protoplasma*, **19**, 1–25.

Schreiber, H. 1934. Strahlenbiologische Untersuchungen besonders im ultravioletten Spektralbezirk an *Saccharomyces turbidans* Hansen. *Strahlentherapie*, **49**, 541–595.

Teindl-Czech, L. 1937. Über die Generationsdauer der Hefe in Abhängigkeit von Strahlenwirkungen. *Protoplasma*, **27**, 313–338.

Warburg, C. 1928. The chemical constitution of respiration ferment. *Science*, **68**, 437–443.

THE CORRODED BRONZE OF CORINTH

EARLE R. CALEY

Department of Chemistry, Princeton University

(Communicated by Hugh S. Taylor)

Abstract

In an effort to determine why bronze objects found at the site of ancient Corinth are almost invariably in a very poor state of preservation chemical analyses were made of several typical objects. Two especially significant facts were thus discovered. One was that the metal of all such objects originally contained a high proportion of tin and little or no lead, a type of metal likely to become highly corroded under natural conditions. The other was that basic cupric chloride was present in noticeable quantity in the corrosion products of these objects, an indication that the soil conditions at the site were unfavorable for the preservation of bronze. That the water and soil at Corinth contain an unusually high proportion of soluble chlorides is indicated by the presence of an unusually high proportion of chloride ion in the water of the Fountain Peirene. Experiments in which a clean fragment of ancient Corinthian bronze was exposed to the action of chloride solutions under various conditions confirmed the hypothesis that soluble chlorides, as well as the composition of the metal, were responsible for the very poor state of preservation of the objects.

In order to obtain information on the mechanism of the corrosive action, a detailed microscopic examination was made of the metallic core and patina of one of the objects. Photomicrographs of typical structures were taken, and some of these are shown on the Plates.

From the various observed facts and the general electrolytic theory of corrosion a detailed explanation is given of the probable course of the chemical changes which occurred during the corrosion of bronze at Corinth. This explanation fully accounts for the formation of the various compounds found in the corrosion products.

In conclusion a critical examination is made of the statement of Pausanias about the treatment of the bronze of Corinth in the water of the Fountain Peirene. It is shown that chemical and metallurgical facts make possible the definite rejection of certain previous interpretations or translations of this statement. For example, it is impossible that the treatment had as its purpose the hardening of the metal. What appears likely is that the water was employed as a cooling medium in the working of high-tin bronze and that the successful working of this particular alloy at Corinth was wrongly ascribed to the peculiar character of the water of the Fountain Peirene rather than to the particular course of heat treatment to which the bronze was subjected.

Introductory Remarks

The unusually poor state of preservation of bronze objects recovered at the site of ancient Corinth has been frequently noted by archæologists who have excavated there. Out of curiosity an investigation was undertaken for the purpose of finding some explanation for the poor state of preservation of such objects, and it is believed that an adequate explanation has been found.

The facts supporting this explanation and the explanation itself constitute the principal substance of this paper. However, various new specific facts or general conclusions of archæological or chemical interest, not directly related to the main object of the investigation, were uncovered or became apparent in the course of the work. The most important or interesting of these are also reported. In addition, a critical interpretation is given of the somewhat obscure remark of Pausanias concerning the treatment of bronze at Corinth with the water of the Fountain Peirene.

Chemical Analyses of Metals and Corrosion Products

In order to determine whether or not some inherent chemical peculiarity of the Corinthian bronzes was responsible for their poor state of preservation, a series of typical objects in various states of preservation was analyzed. The results of these analyses are listed successively in accordance with the ages of the objects analyzed. Significant facts revealed by each analysis are discussed immediately after the tabulation of each set of analytical results, but general conclusions germane to the main thesis of this paper are given only after the presentation and discussion of all the individual analyses.

Middle Helladic Ring.—This small bronze ring was so brittle and corroded that the excavators on removing it from the ground had accidentally broken it into several fragments. These fragments could be broken easily into smaller fragments. From the freshly broken ends of such fragments it was evident that each contained a central core of distinctly metallic appearance, and that this core was surrounded by a thick, dense, reddish layer of corrosion products, which in turn was surrounded by a thin exterior coating of clay soil admixed with small amounts of greenish corrosion products. On qualitative examination the thick reddish layer was found to consist almost entirely of cuprous oxide. As will be shown quantitatively from the examination of some of the other objects, this particular compound is the principal product formed in the corrosion of bronze at Corinth. Before analysis all the material surrounding the central cores of the fragments was removed mechanically. The isolated cores were metallic and homogeneous in appearance, though their brittleness and the color of filings detached from them indicated that considerable cuprous oxide was present in the metal as an intergranular corrosion product. No attempt was made to separate this compound from the metal before analysis since it seemed impossible to do so without altering the composition of the metal itself. The following analytical results therefore represent the composition of the metallic core of the object freed from corrosion products external to its surface:

Cu	80.98
Sn	13.90
Pb	.57
Fe	.05
Ni	.04
Co, Zn, Ag, Au	nil
As	.51
S	nil
Total	96.05
O (By Difference)	3.95
	100.00

These results show the metal of this ring to be a simple high-tin bronze containing such a small proportion of lead that this metal must be regarded as an impurity and not as an intentional component of the alloy. Iron, nickel, and arsenic, the other impurities found, are the ones often encountered in ancient bronze. Only the arsenic is present in proportion somewhat greater than usual. Sulfur, an impurity that might lead to rapid corrosion, is absent.

Geometric Bowl.—Only a few, small, completely corroded fragments of this object were available for analysis, and these fragments were so thin and the substance of them so mixed with adherent soil that it was not possible to obtain a sample of the corrosion products sufficiently homogeneous and representative to justify a quantitative analysis. Therefore, only a qualitative analysis was made. The following results were obtained:

Present in Large Proportion	Cu
Present in Moderate Proportion	Sn, O
Present in Small Proportion	Pb, Fe, Cl
Present in Very Small Proportion	Zn
Absent	Ni, Co, Ag, Au

Elements such as aluminum, which were found to be present, but which obviously occurred in the contaminating soil and not in the corrosion products of the metal, are purposely omitted from this tabulation. According to these results the original metal was an ordinary bronze containing some of the usual impurities. However, the presence of metallic chlorides in the corrosion products of this alloy is not so usual, and their presence is of special significance for explaining the corrosion of buried bronze at Corinth.

Sixth Century Bowl.—For the same reasons as given for the preceding object no quantitative analysis was attempted. The results of the qualitative examination were as follows:

Present in Large Proportion	Cu
Present in Moderate Proportion	Sn, O
Present in Small Proportion	Fe, Cl
Absent	Pb, Ni, Co, Zn, Ag, Au

According to these results the original metal was a bronze that contained no lead and was unusually free from some of the other metallic impurities often found in ancient bronze. The absence of nickel in both this corroded object and in the preceding one is worthy of particular remark since nickel is almost invariably present as an impurity in unaltered bronze found at Greek sites. It is to be especially noted that metallic chlorides were also present in the corrosion products of the metal of this object.

Sixth or Fifth Century Strigil Blade.—This thin blade was completely corroded and had been broken into many small fragments. Except for some adherent soil and for a very thin layer of a green copper corrosion product, apparently basic cupric chloride, on their outside surfaces, these fragments were composed of a brownish-red substance that appeared to consist mostly of cuprous oxide. Since the material was sufficient in amount, and representative samples reasonably free from adherent soil could be prepared, a complete quantitative analysis was made. The results were as follows:

Cu	75.83	SiO_2	1.44
Sn	7.98	CO_2	tr.
Pb	.07	Combined H_2O	pr.
Fe	.89		
Ni	.01		87.08
Co, Zn, Ag, Au	nil	O and H_2O (By Difference)	12.92
As	.08		
S	.05		100.00
Cl	.73		

The original metal of this corroded object, like that of the two preceding objects, was evidently a simple bronze, though one that contained more of the impurities usually found in ancient bronze. From the ratio of tin to copper given by this analysis, the alloy of which the object was originally composed would appear to have contained about ten per cent of tin. However, it is very likely that the percentage of tin was considerably higher than this, since, as will be clearly shown from analyses of the next object, the ratio of tin to copper tends to be lower in the corrosion products of an ancient bronze than in the unaltered metal. Hence this strigil blade was probably composed of high-tin bronze.

The proportion of sulfur found by this analysis is probably too small to be of any significance for explaining the complete corrosion of the object. Similar low proportions have been frequently found in well preserved ancient bronzes. On the other hand, the proportion of chlorine, which is much higher, is undoubtedly of some significance for explaining the extent to which the object was corroded. The marked proportion of silica found on analysis may have been present in the form of soil on the exterior of the fragments that were analyzed, though because of the care taken to remove all visible soil, it seems much more likely that this silica was actually present within the mass of the fragments as a deposit from subterranean water containing dissolved silica. That the subterranean water at Corinth contains much dissolved silica is certain from the composition of the water of the Fountain Peirene, an analysis of which is given in the next part of this paper. The presence of only a trace of carbonate in the material of this object is perhaps surprising in view of the general belief that basic copper carbonate is an invariable product of the corrosion of ancient bronze.

The various compounds present in the mixture of bronze corrosion products of which this object is composed may be regarded as true minerals formed by natural processes. The identity of the principal compounds or minerals which occur in this mixture of corrosion products may be established with reasonable certainty from their physical and chemical properties, and the approximate proportions in which they occur may apparently be calculated with equal certainty from the results of the analysis. If all the chlorine is assumed to be present in combination as the basic cupric chloride, atacamite, with the formula $3Cu(OH)_2 \cdot CuCl_2$, and all the copper not required for combination in this compound is assumed to be in the form of cuprous oxide, or cuprite, and the tin and iron are taken to be present in their highest states of oxidation wholly as anhydrous oxides, then the principal components of these corrosion products, including also the silica, are as follows:

Cu_2O	82.4
$3Cu(OH)_2 \cdot CuCl_2$	4.4
SnO_2	10.1
Fe_2O_3	1.3
SiO_2	1.4
	99.6

Whether the oxidized iron is wholly present as the anhydrous oxide, hematite, or in part as one of the hydrated forms of this oxide, limonite for example, is not quite certain. If present in large part as one of the hydrated forms, which is not unlikely, then the summation of the principal components would be even closer to one hundred per cent. At any rate the very satisfactory summation of the calculated proportions of the principal components of the mixture of corrosion products is at least a good indication of the essential correctness of the identifications and of the method of calculation. Though no attention was paid in this calculation to the various impurities, such as arsenic and lead, it is clear from their small proportion that in whatever form of combination these elements occur in the mixture of corrosion products no significant change in the proportions of the principal components of the mixture could result if they were taken into account. No attempt was made to take them into account because the determination of the exact form of their combinations in such a complex and preponderant mass of other compounds is a problem of great experimental difficulty and uncertainty.

Fifth Century Strigil Handle.—This last object was the most interesting and instructive of the series because it contained a metallic core surrounded by a thick, compact patina which could be sharply separated mechanically from the core, thus making possible a study of the chemical composition of both the metal and its corrosion products. Moreover, as will appear in another part of this paper, the peculiar condition of the object made possible a thorough study of the microscopic structure of both patina and metal, and, more important still, of the region of transition between the two.

This strigil handle, almost complete, was 8.5 cm. long and had a nearly uniform width of 1.5 cm., except near the end where the blade had been attached. There the width was 2.0 cm. It had a nearly uniform thickness of 0.5 cm. except at the blade end where the thickness was only 0.2 to 0.3 cm. The metallic core was continuous throughout the straight part of the handle, and the layer of patina surrounding this core had a nearly uniform thickness which varied only between 0.15 cm. and 0.20 cm. The exterior of the object was nearly free from adherent soil and was of a dull green color. For the most part the very thin coating of copper corrosion products which constituted the outermost layer of the patina consisted of the basic cupric chloride, atacamite, though a small proportion of basic cupric carbonate, or malachite, was also present, and in one isolated spot a few minute crystals of the other natural basic cupric carbonate, azurite, were observed. Between this outermost layer and the metallic core, the patina was largely composed of a very hard, compact, brownish or greyish purple material arranged in a great number of more or less distinct thin layers around the core. A very few thin layers or minute patches of a greyish or greenish white material were irregularly distributed in the dark ground mass of the patina. The purplish constituent of the patina was largely cuprous oxide, and the whitish material appeared to be largely stannic oxide, though some of it apparently contained considerable basic cupric chloride.

For the analyses, a section 3 cm. in length was selected from the middle of the object. The outermost surface of the patina was removed with a file and discarded in order to make sure that no soil or other foreign matter was included in the sample of patina taken for analysis. By means of a clean file the patina was then systematically filed down to within a very short distance of the metallic core, special care being taken, however, to avoid including any filings of metal. The fine powder so obtained had a brownish red color, very different from that of the compact patina. After treating this powder with a strong magnet in order to remove minute particles of iron that had been detached from the file, and after mixing it thoroughly, the powder was analyzed directly without subjecting it to any drying treatment. In order to obtain samples of the metal, the remainder of the patina was filed away from the metallic core and discarded. Though the metal so isolated had a bright, brassy lustre, its brittleness and the color of filings detached from it indicated that some cuprous oxide was still present within its mass as an intergranular corrosion product. No attempt was made to separate this compound from the metal before analysis because it seemed impossible to do so without altering the composition of the metal itself. The cleaned metallic core, which was very hard, was then divided into several samples of suitable size. The following results were obtained from the analysis of these samples:

Cu	82.23	Co, Zn, Ag, Au, As, S	nil
Sn	13.86		
Pb	.02	Total	96.34
Fe	.21	O (By Difference)	3.66
Ni	.02		
			100.00

These results show the metal to be a simple high-tin bronze unusually free from the various impurities often found in ancient bronze. Even the lead content is abnormally low.

Analysis of the sample of patina yielded the following results:

Cu	76.54	CO_2	tr.
Sn	7.34	Combined H_2O	pr.
Pb	.02		
Fe	.27	Total	84.88
Ni, Co, Zn, Ag, Au, As, S	nil	O and H_2O (By Difference)	15.12
Cl	.30		
SiO_2	.41		100.00

A comparison of these analytical results with those obtained on the preceding object will show at once that this patina, formed by the partial corrosion of a particular kind of bronze object, is very similar in composition to the material resulting from the complete corrosion of a similar object. There are, of course, certain minor qualitative and quantitative differences. Thus, the small proportions of arsenic and sulfur present in the one are absent from the other, but such differences arise from insignificant differences in the composition of the metals, not from any essential difference in the processes of corrosion. The presence of both chloride and silica, and the absence of any marked proportion of carbonate, in both the patina and the corrosion products of the strigil blade are especially to be noted. The presence of silica in this carefully prepared sample of patina appears to confirm the previous tentative conclusion that the silica found upon the analysis of the corrosion products of the blade arose not from extraneous soil but from free silica contained within the body of the corrosion products. Not only does the similarity of the two sets of analytical results indicate that the same compounds or minerals are present in the patina and in the corrosion products of the blade, but this similarity also indicates that they are present in approximately the same proportions. By means of the method of calculation outlined before, it may be estimated that the patina contains 85.0 per cent of cuprous oxide as against 82.4 per cent for the corroded strigil blade, 9.3 per cent of stannic oxide as against 10.1 per cent for the corroded blade, and 1.8 per cent of basic cupric chloride as against 4.4 per cent for the blade. That the same end products were formed in about the same proportions is at least an indication that corrosion took place under the same conditions and that the chemical processes involved were the same. It should perhaps be noted in passing, that, as far as the author has been able to determine, this analysis of the patina of a Greek strigil handle is the first complete quantitative analysis of the patina of any sort of a Greek bronze object.

Some interesting conclusions follow from a comparison of the composition of the patina of this object with the composition of the bronze from which it was formed. In the first place the proportion of tin to copper in the patina is much lower than in the alloy itself. According to the results of the analyses the ratio of tin to copper in the patina is only about one to ten, whereas in the alloy it is about one to six. Various explanations for this marked difference are possible, but, regardless of the particular way in which it came about, any calculation of the ratio of tin and copper in the unaltered alloy of this object from the ratio found in the patina, on the assumption that this remained constant, would have led to a very wrong conclusion, and wrong conclusions would be obtained from similar calculations of the composition of the unaltered metal of other objects of like composition. Yet, on the basis of this very assumption, various chemists have confidently calculated the original composition of ancient bronzes from the results of analyses of their corrosion products. In view of what has been shown in regard to the ratios of tin and copper in the patina and metal of this strigil handle, the correctness of compositions so calculated may seriously be doubted. However, from a number of analyses of different types of bronzes and of their corrosion products formed under different conditions, it may be possible to derive a set of correction factors that would make possible the close calculation of the original proportions of the principal components of an ancient bronze from the results of an analysis of its corrosion products. Whether the original proportions of the minor components or impurities can ever be so calculated is doubtful. In support of this statement it will be noted that no nickel was detected in the patina of this strigil blade, though 0.02 per cent, a small but easily determinable proportion, was found in the metal. This difference perhaps explains why no nickel was found in the completely corroded remains of the two bronze bowls, and why only 0.01 per cent was found in the corroded strigil blade. Apparently the oxidation products of nickel are leached out of the corrosion products of buried bronze by the action of subterranean water. On the other hand, as far as the results of the present analyses are indicative, lead appears to remain in the corrosion products. It will be noted that the proportion of lead is the same in the patina of this strigil handle as in the metallic core. Neither is any loss of iron apparent. Indeed, according to the actual analytical results, the proportion of iron is higher in the patina of this strigil handle than in the metal itself. Iron may possibly occur in higher proportion in the corrosion products of a bronze than in the unaltered metal because of addition of iron from the soil or from subterranean water. Whether these present findings concerning the relative proportions of certain elements in unaltered bronze and its corrosion products are really of general validity can only be determined by complete analyses of many more ancient bronzes together with similar analyses of the corrosion products of these bronzes.

Partly for the sake of gathering together in one place all the known analyses of Corinthian bronze objects other than coins, and partly because of its value as an independent confirmation of the present results concerning the type of alloy employed in the fabrication of the bronze objects found at Corinth, one other analysis is given here, that of the metal of a Corinthian bowl analyzed by Rammelsberg.[1] After an extensive search through journals and books this was the only previous analysis of such a Corinthian object that could be located. Unfortunately, the state of preservation of this bowl is not mentioned nor is it dated, though a brief description given of its form indicates that it was manufactured in the Greek period. The bronze of this bowl is stated to have the following composition:

Cu	86.87
Sn	11.91
Pb	.72
Fe	.25
Ni	.25
Total	100.00

The most likely explanation for the perfect summation of these analytical results is that copper or some other element was estimated by difference, or that the determined percentages of all the elements were recalculated to a basis of one hundred per cent. Use of the second method for arriving at the final results would be an indication that the object was more or less corroded. However this may be, the main point of interest in these results is that they show the metal of this bowl to be a high-tin bronze of the same type as the metal of the other Corinthian objects here studied as to quantitative composition.

In view of the probable use also of high-tin bronze containing very little lead in the earliest, or fourth century, Greek bronze coins struck at Corinth,[2] all the evidence at present available appears to show that bronze of this one particular type was very commonly, perhaps invariably, used for the fabrication of all sorts of bronze objects at Corinth over a very long period. Whether the manufacture and use of such a bronze was a traditional metallurgical practice peculiar to Corinth during this time cannot be determined yet because of lack of sufficient data on the composition of bronze from other ancient centers of bronze manufacture.

---

[1] Cited by Pabst, *Archäologische Zeitung,* 36, 164 (1878).

[2] Caley, "The Composition of Ancient Greek Bronze Coins," Amer. Philos. Soc., Mem. XI, Philadelphia, 1939, p. 67.

RELATIONSHIP OF COMPOSITION TO STATE OF PRESERVATION

From a study of the relationship of composition to state of preservation in a considerable number of ancient Greek bronze coins,[3] it is known that high-tin bronze containing very little lead, that is, bronze of the type composing these Corinthian objects, is one of the types of bronze most likely to become highly corroded under natural conditions. A specific indication of the importance of original composition for the preservation of bronze and copper objects at Corinth is given by the composition and state of preservation of an Early Helladic needle from this site. Though older by at least a thousand years than the earliest of the bronze objects here investigated, and older by about two thousand years than the latest of these objects, this needle was, nevertheless, in a better state of preservation than any of them. The chief reason for its excellent condition became evident from the results obtained on analysis which showed this object to be composed, not of bronze, but of arsenical copper, a metal rather resistant to the action of natural agents of corrosion. Undoubtedly, therefore, the present poor state of preservation of these Corinthian bronzes may be ascribed, in part at least, to their original chemical composition.

However, their present condition cannot be wholly explained on the basis of original composition since high-tin bronzes containing little or no lead have been found in a good state of preservation at some sites. That the bronzes buried at Corinth were subjected to ground conditions not favorable to their preservation is indicated by the presence of basic cupric chloride in their corrosion products. This compound was probably formed as the result of reactions between the copper in the alloy and subterranean water containing chlorides, or as a result of reactions between this metal and moist earth containing chlorides deposited from such water. That bronze is corroded at an appreciable rate by water containing chlorides in sufficient concentration has been repeatedly demonstrated by experiments on the corrodibility of bronze. The composition of the subterranean water at Corinth was therefore investigated in order to determine to what extent the poor state of preservation of Corinthian bronzes could be attributed to the ground conditions at that site.

[3] Caley, *op. cit.*, pp. 180–185.

Composition of the Subterranean Water at Corinth

The principal source of water at the site of Corinth is the remarkable spring known as the Fountain Peirene. That this spring was also the principal source of water for the ancient city is certain from an abundance of archæological and historical evidence. Though the nature of the rock formation in the neighborhood indicates that a natural spring existed here even in very remote times, the rate of flow of this original spring was evidently much increased artificially by the ancients who cut large reservoirs and supply tunnels in the rock so that the subterranean water from a considerable space in the high ground immediately above the city was collected and caused to flow out into convenient draw basins. These ancient reservoirs and tunnels are still in such an excellent state of preservation that a copious flow of water, amounting to about 25,000 liters an hour on the average, is even today obtained from this spring. In fact, there seems to be no reason for believing that the rate of flow differs materially now from what it was in ancient times. Nor is there any reason, from geological or other considerations, for believing that the chemical composition of the water differs materially now from what it was in ancient times, or, indeed, from what it was at any time during the long intervening period. The composition of the water of the Fountain Peirene is obviously representative of the composition of the whole body of subterranean water above the ancient site, and it is very probably representative also of the composition of the subterranean water of the site itself, not only because these waters are in the same small locality in adjacent rock strata but especially because any subterranean water draining into the site of the ancient city must necessarily come from the high ground above it. The present chemical composition of the water of the Fountain Peirene should therefore be representative of that of the water which came into contact with the ancient bronze objects now found at Corinth. No direct sample of the subterranean water of the site could be taken at the time of this investigation because of dry conditions. However, the composition of the water of the spring is probably more representative of the general composition of the subterranean water reaching the site sporadically or otherwise over the centuries than is the composition of any small local sample of subterranean water taken at one particular time at the site. Therefore, a complete mineral analysis of a large typical sample of the water of the spring was made as a means of obtaining the desired information about the normal composition of this subterranean water. The following results were obtained:

Ion or Compound Present	Amount Present in Parts per Million (Milligrams per Liter)
Calcium (Ca^{++})	62
Magnesium (Mg^{++})	57
Potassium (K^{+})	48
Sodium (Na^{+})	94
Iron (Fe^{++})	0.2
Bicarbonate (HCO_3^{-})	550
Chloride (Cl^{-})	106
Sulfate (SO_4^{--})	48
Dissolved or Colloidal Silica (SiO_2)	63
Nitrogen Compounds and Organic Matter	Present but not Estimated
Total Dissolved Matter	1030

These results show this water to be essentially a solution of the bicarbonates, chlorides, and sulfates of calcium, magnesium, potassium, and sodium containing in addition considerable dissolved or colloidal silica and also small amounts of organic substances and nitrogen compounds derived from these substances. This water is interesting from a general chemical and mineralogical standpoint not only because these dissolved substances as a whole are present in very high concentration but because certain of these substances are present in unusually high concentration. Most interesting and significant from the special standpoint of this particular investigation, however, is the unusually high chloride content of the water. This chloride content is entirely sufficient to account for the formation of the small amounts of basic cupric chloride found in the corrosion products of the ancient bronze objects from Corinth. It also helps to account for the degree to which these objects are corroded, as will be shown in detail shortly from both an experimental and theoretical viewpoint.

Since there are no salt deposits near Corinth, the chloride in this spring water apparently originated from salt spray whipped up by wind from the nearby Corinthian Gulf, or from the not very distant Saronic Gulf, this spray being washed directly into the ground during rainy weather, or else being washed down by subsequent rains after drying on the surface of the ground, or after drying in the air and being deposited on the ground as a dry salt mixture. That sodium chloride is transported considerable distances inland from bodies of salt water in just such ways has in fact been conclusively demonstrated by several systematic investigations. Especially pertinent here is the investigation conducted continuously over a two-year period by Bordas and Desfemmes[4] on the French Mediterranean coast at Cette and at neighboring inland localities. These investigators analyzed samples of dry surface dust and of rain water collected at different distances from the sea and at different altitudes. Considerable sodium chloride was found in dust collected 10 kilometers inland from Cette, and salt was even detected in dust collected 50 kilometers from the sea. Rain water collected at Vasquières, 35 kilometers inland from Cette, was found to contain on the average about 2.7 milligrams of sodium chloride per liter, or about a twentieth of that contained in the rain water at Cette immediately on the coast. The largest quantities of sodium chloride were found after a severe dust storm when the salt deposited on the surface of the ground amounted to as much as 8.4 grams per square meter. The smallest quantities were found after several days of consecutive rain, obviously because salt deposited on the ground as dust had been washed into the ground or had been removed in the water draining off the ground. In general, this investigation of Bordas and Desfemmes shows that much more sodium chloride is transported inland in the form of dust in fair weather than is transported in rain during stormy weather. It would appear very likely, therefore, that sodium chloride would tend to be deposited on the surface of the ground in much greater amounts at a comparatively dry locality such as Corinth than at a damp locality situated a similar distance from salt water. It also follows that rain falling on the site of ancient Corinth after a prolonged dry spell, on seeping into the ground, probably brought water containing much chloride into contact with the buried bronze objects, the action of such water in furthering the corrosion of these objects being supplementary to, perhaps often independent of, that of the subterranean water draining down into the site from higher ground.

The abnormally high silica content of this spring water is also significant from the standpoint of the present investigation. This silica content is entirely sufficient to account for the amounts of silica found in the corrosion products of the ancient bronze objects. Possibly silica is often deposited by subterranean water within the corrosion products or patina of buried bronze. This would explain the unusual hardness of some bronze patinas, a hardness that is not easily explicable from the moderate hardness of the copper minerals that are the principal components of such patinas. This is a subject that could well be investigated further.

---

[4] *Compt. rend.*, **185**, 603–605 (1927).

EXPERIMENTAL CORROSION OF CORINTHIAN BRONZE

In order to show by direct experiment that water containing chloride ion was in part responsible for the poor state of preservation of Corinthian bronze objects, and in order to determine what products were formed by the action of such water on Corinthian bronze under various conditions, a number of experiments were performed on the action of dilute sodium chloride solutions on a clean sample of this bronze. The sample used in the experiments was a selected part of the metallic core of the strigil handle before described. Though this bronze sample was definitely corroded by a sodium chloride solution having the same chloride-ion concentration as that of the water of the Fountain Peirene, and though the apparent effects of this corrosive action were exactly analogous to those caused by more concentrated solutions of sodium chloride, the rate of corrosion by this solution was too slow for convenient experimental purposes, especially since the amounts of corrosion products formed in a reasonable time were too small for certain identification. However, it was found that a one-tenth of one per cent solution of sodium chloride, a solution only about six times more concentrated in chloride ion than the water of this spring, did corrode the bronze sufficiently rapidly and did yield sufficient corrosion products in a reasonably short time. The following experiments illustrate the most significant results obtained in the various experiments with this solution.

Experiment I

The clean bright sample of bronze was alternately dipped into the sodium chloride solution and exposed to air several times a minute by means of a mechanical contrivance. Signs of active corrosion were apparent in less than an hour. In a few hours the surface of the metal was completely covered with an adherent red corrosion product. The solution remained clear. This red corrosion product was mostly cuprous oxide, but it contained an insoluble tin compound, and may have contained metallic copper. Both copper and tin were present in the clear solution.

Experiment II

The cleaned bronze sample was placed in a bottle partly filled with the solution, and the bottle was then stoppered. Within a few hours signs of active corrosion were apparent. At the end of 24 hours the metal was covered with a loose red corrosion product. A white corrosion product was present as a fine suspension in the solution around the bronze sample, and as a fine powder on the bottom of the bottle immediately around the sample. The red corrosion product appeared to have the same composition as that obtained in the first experiment. The white product was mostly cuprous chloride, but it also contained an insoluble tin compound. Again both copper and tin were present in the solution.

Experiment III

The cleaned bronze sample was placed in a bottle completely filled with the solution from which dissolved air had been previously removed by boiling. The bottle was then tightly stoppered. Again signs of active corrosion were apparent in a few hours. At the end of 24 hours a very small amount of the red corrosion product was present on the surface of the metal, and a very small amount of the white corrosion product was present as a suspension in the solution around the sample. No further corrosive action was evident after continuing the experiment for a number of days. The amounts of corrosion products obtained were too small for certain analytical identification, but they appeared to be the same as those produced in the preceding experiment.

Continued corrosion of the bronze sample under the conditions of the second experiment for a number of days was found to result in the formation of small amounts of basic cupric chloride in addition to cuprous oxide and cuprous chloride. In none of the experiments were corrosion products of metals other than tin and copper detected, probably because the amounts of such products formed in the comparatively short time allowed for even the longest experiment were too small to be detected by the analytical methods used.

These experiments clearly demonstrate how readily Corinthian bronze is corroded by very dilute sodium chloride solutions. The experiments are of additional significance because they show that the copper corrosion products experimentally produced by the action of very dilute sodium chloride solutions on Corinthian bronze are, in part at least, identical with those actually found on the ancient corroded bronze. It seems particularly significant that cuprous oxide, the principal compound found in the corrosion products of the ancient bronze, may also be produced, under certain experimental conditions, as the principal compound resulting from the corrosive action of very dilute sodium chloride solutions. Moreover, basic cupric chloride, which occurs in small proportion in the corrosion products of Corinthian bronze objects, may also be produced experimentally in similar small proportion. On the other hand, cuprous chloride, which does not occur in any noticeable proportion among the corrosion products of the ancient objects, may be readily produced under certain experimental conditions as the principal corrosion product. But this sharp difference in the nature of the experimental and natural corrosion products may be easily explained. In the first place, the natural corrosion of the bronze may have occurred in the presence of an abundant supply of oxygen, as by the alternate action of air and water, a condition simulated in the first of the experiments described, so that cuprous chloride was not formed as a natural corrosion product, except perhaps momentarily. Moreover, cuprous chloride is such an unstable compound in the presence of oxygen and moisture that even if it had been formed originally in quantity as the result of corrosion by slightly saline water in the absence of a free supply of oxygen, it would have been largely or entirely transformed during the centuries of burial when oxygen and moisture were undoubtedly present in sufficient amount at many times. Even if small amounts of cuprous chloride had been present, as a result of recent corrosive action, in the corrosion products of the objects at the time of their excavation, it is unikely that this compound would have remained unaltered during he several years that elapsed between the time at which most of hese objects were taken from the ground and the time at which they were subjected to chemical analysis. An apparent exception to this general absence of cuprous chloride in the corrosion products of these ancient objects was found, however, during the course of a systematic microscopic examination of a freshly cut section of the strigil handle. The evidence for the presence of cuprous chloride in this object is given in the next part of this paper where the microscopic structure of this object is described in detail. The essential identity of the copper corrosion products of ancient Corinthian bronze with those experimentally produced by the action of very dilute chloride solutions appears to indicate, therefore, that these natural corrosion products were also formed by the action of water containing chloride.

From the composition of the subterranean water at Corinth, the results of the experiments, and the nature of the corrosion products, it may be safely concluded that the present poor state of preservation of the bronze objects found at Corinth is not to be ascribed wholly to the composition of the bronze, but is to be ascribed also to the peculiar ground conditions at that site. Indeed, since bronzes of the same composition as those of Corinth have been found in good condition at some sites, these ground conditions may well have been a much more important cause of corrosion than the composition of the metal. The two causes acting together certainly account adequately for the poor state of preservation of the bronze objects now found at Corinth.

Microscopic Examination of the Metallic Core and Patina of a Corroded Corinthian Bronze Object

In order to see in detail what physical changes occur in the structure of Corinthian bronze as a consequence of the long continued corrosive action of the local water, and in order to investigate further the nature of the corrosion products of this bronze with the end in view of obtaining more information about the chemical changes involved in the formation of these products, one of the specimens of corroded bronze was subjected to a thorough microscopic examination.[5] The specimen selected for this examination was part of the strigil handle described previously. Because of the deep but still incomplete corrosion of this bronze handle, the hardness and solidity of the layers of corrosion products, and the presence of a sharp transition zone between these corrosion products and the metallic core, this particular object was almost ideal for the purpose. The section of the strigil handle selected for examination was taken from near the end where the blade had been attached. At the blade end of this selected section the metallic core was much thinner and the layer of corrosion products much thicker than at the other end where the core and the layer of corrosion products were of the same thickness as in the remainder of the handle. Consequently, one end of the section represented a more advanced stage of corrosion than the other, and by examining each end separately the effects connected with two successive stages in the progress of the corrosion could be separately studied.

---

[5] Nearly all the experimental work described in this section of the paper was performed by Mr. C. R. Stevens, a former student in chemistry at Princeton Univer-

In order to handle the specimen conveniently, both in preparing it for examination and in examining it with the metallographic microscope, it was mounted lengthwise within a small brass cylinder by means of an alloy of low melting point. This method of mounting was found to be especially satisfactory for a specimen of this nature since it was so firmly held in the imbedding medium that no flakes of the brittle corrosion products were detached on grinding or polishing as they undoubtedly would have been if the specimen had been ground and polished by itself. The imbedded specimen, the imbedding medium, and the brass cylinder were filed and ground down at a right angle to the axis of the cylinder so as to form smooth plane surfaces at each end, in the middle of which were complete cross sections of the metallic core and its surrounding layer of corrosion products. The grinding was performed by rubbing the mounted specimen on emery papers of successive degrees of fineness fastened to smooth wooden blocks. After the end surfaces of the imbedded specimen and its mount had been smoothly ground they were highly polished with levigated alumina of three successive grades of fineness by rubbing these surfaces on moist cloths impregnated with this polishing powder and stretched on smooth

sity, as part of a senior thesis on the metallography of ancient bronze. Many of the interpretations given here of this experimental work are also to be credited to him. The author is especially indebted to Mr. Stevens for his care and skill in preparing the excellent photomicrographs which are reproduced on the plates.

wooden blocks. In accordance with the kind of microstructure to be revealed, the polished surfaces of the specimen were then etched with some one of the reagents commonly used in treating bronze for metallographic examination. Of these various reagents a ten per cent solution of ammonium persulfate was found to be the most generally useful. The process of polishing and the process of etching usually had to be repeated a few times for each different part of the specimen being examined before entirely satisfactory views of the local microstructure were obtained since the choice of reagent and the method of using it had to be experimentally determined for nearly each different part of the heterogeneous specimen. Photographs were taken of the most interesting and significant structures seen under the microscope, some fifty photomicrographs being taken in all. Of these numerous photomicrographs, ten representative ones are reproduced on the plates.

The upper figure of Pl. I is a view of the best preserved part of the metallic core of this section of the strigil handle. From this figure it will be seen that the physical structure of the uncorroded metal of this object consists for the most part of homogenized and twinned crystals or grains of variable but generally small size. No dendritic crystals are to be seen. This type of structure indicates that the object was fabricated by some mechanical means, not by casting. Very likely it was formed by hammering, the handle and the blade of the strigil probably being hammered out from a single piece of bronze of suitable size. The considerable variation in the size of the metal grains, the presence of annealing bands on the faces of many of these grains, as well as the chemical composition of the metal as determined by analysis, show beyond a doubt that this object was subjected to elevated temperatures in the course of its fabrication. Since the bronze contained nearly 14 per cent tin, the use of elevated temperatures in forging the metal was almost necessary because bronze containing such a high proportion of tin is too hard , stiff, and brittle to be shaped readily by hammering it while it is cold. In fact, bronze of this composition cannot be successfully forged at temperature below 590° C. unless it has been given a special annealing treatment beforehand. By carefully annealing it for a long time or by heating it to above 650° C. and quenching it, such bronze may be worked while it is cold but this is a more troublesome process than hot forging. It seems likely that the simpler process would have been the one followed by the ancient metallurgists at Corinth, though from the appearance of the microstructure the use of the other is by no means excluded. The presence of more than one system of annealing bands on some of the grains seems to indicate that the object was repeatedly heated and hammered in the course of its fabrication. Probably the metal was heated to a bright red heat, hammered until it was too cold to be shaped readily, then heated again to a bright red heat and again hammered, the alternate heatings and hammerings being repeated until the object had the desired shape. The numerous slip lines to be seen on some of the grains do indicate, however, that the metal was worked to some extent while it was at much lower temperatures than those suitable for hot forging. On a few grains more than one system of slip lines are visible. The presence of these is not clearly shown in this top figure of Pl. I, but it is very clearly shown in the bottom figure of Pl. I where the magnification is much higher. Possibly the occurrence of these slip lines only indicates that the object had accidentally cooled down below the proper forging temperatures while it was receiving its final shaping, or the occurrence of these lines may indicate that the object was deliberately hammered at lower temperatures after it had been shaped. The effect of such hammering would be to impart additional hardness and stiffness to the thin blade and handle, and in view of the nature of the object this may well have been done intentionally. On the other hand, the presence of these lines may indicate that the object was shaped to some extent by cold working after being quenched from a sufficiently high temperature. Possibly this was done only in the last stages of manufacture when the thin object may have cooled off too rapidly for convenient forging at a high temperature. Whatever may be the exact significance of the presence of these slip lines, it seems certain that the metal was worked to some extent at relatively low temperatures, even though most of the process of manufacture was performed at elevated temperatures. The microstructure of the metal therefore yields considerable information about the method employed in fabricating this particular object. As will appear shortly, some of this information about the method of manufacture is of value for explaining the peculiar pattern of the corrosion within the body of the metal.

The top figure of Pl. I shows also that some corroded metal is present even within the best preserved part of the metallic core of this object. In the photomicrograph the corroded metal appears as irregular black spots and as heavy black irregular lines. It will be observed that a few of the metal grains have been completely corroded, and that considerable intergranular corrosion has occurred, in a few places to such an extent that grains of metal are completely surrounded by corrosion products. A more highly magnified view of the corroded metal in this part of the core is shown in the lower figure of Pl. I. In the top figure of Pl. II is shown at the same high magnification a few grains of metal completely surrounded by intergranular corrosion products. The view in the photomicrograph is, of course, only a cross section. In the solid mass of the metal such grains are probably enclosed in thin shells of corrosion products. This existence of highly corroded metal in the mass of largely uncorroded metal is of importance as showing that the corrosion of the metal object did not proceed uniformly inward from its surface. Instead, the corrosion evidently progressed in a very irregular manner down along the boundaries of the grains, such intergranular corrosion being always far in advance of the general corrosion of the metal. This seems to imply that in the initial stage of corrosion much of the outside surface of the object remained relatively unaltered while in some places the corrosion along grain boundaries extended a considerable distance below the surface.

The lower figure of Pl. II shows an advanced stage in the corrosion of grains of metal. This view was taken near the edge of the metallic core. Here corrosion is not confined to the boundaries but has progressed toward the center of each of the grains. It will be noted that some of the grains appear to be completely corroded, others nearly so, and still others only to a small extent. These differences are probably not so great as they seem because the view in the photomicrograph is that of a cross section in which the grains, or the corroded remains of grains, have been cut through at various distances from their centers. Considerable differences in the extent of the corrosion of the grains undoubtedly exist, however, because of the widely different original sizes of these grains. If the corrosion of the grains beginning at their boundaries is assumed to have taken place toward their centers at a more or less uniform rate, and this assumption is a reasonable one, then the corrosion of small grains would be complete before the corrosion of large grains had advanced very far. It is very interesting to see that the positions of the original boundaries of the metal grains are very clearly marked by the dark appearance of the shells of corrosion products first formed in the grain boundaries. The sizes and shapes of the original grains of metal may thus be easily discerned in spite of their extensive destruction through corrosion.

From what has now been shown about the way in which this bronze corroded, it is clear that an intergranular corrosion of the metal preceded its general corrosion. In other words, corrosion began at the boundaries of sound grains of metal and progressively advanced into these grains from all directions so that they became more and more corroded. This process finally resulted in the complete corrosion of the metal.

That this progressive corrosion of the grains of the bronze did not proceed uniformly throughout the body of the metal, nor uniformly downward from its surface, is shown in the top figure of Pl. III, which is a view at low magnification of the general pattern of the intergranular and granular corrosion within the metallic core of the object. The places where the metal has been extensively corroded show up as dark markings or streaks, the places where it has been but slightly corroded as light markings or streaks. On the plane surface of the metal seen in the photomicrograph the corroded metal appears to be in the form of thin broken bands which alternate with similar bands of uncorroded metal. A more highly magnified view of this apparently banded structure is shown in the bottom figure of Pl. III. Since these alternate bands were found still to persist in the same relative positions when the metal surface was ground down further to show other cross sections, it is clear that the uncorroded and corroded metal in the object really exist in alternate layers. Moreover, there are two systems of these alternate layers, and they cross each other. This is very clearly shown at the left of the top figure of Pl. V, where the metal near the patina is to be seen in cross section at low magnification. Since all these views are of cross sections perpendicular to the axis of the object, the alternate layers of corroded and uncorroded metal are evidently parallel to this axis. They are not, however, parallel to the sides of the object. This occurrence of alternate layers of uncorroded and corroded metal in the core of this object appears to constitute a type of microstructure that has never been observed before in ancient corroded bronze. The corrosion of the bronze throughout numerous thin layers in advance of its general corrosion is probably an indication of the existence of some corresponding layer-like heterogeneity in the metal before corrosion started. This heterogeneity may have consisted of layers of non-metallic impurities between layers of metal free from such impurities, of alternate layers of metal containing different proportions of these impurities, or of alternate layers of metal containing different proportions of tin and copper with or without differences in the distribution of impurities. Such layers might conceivably have been formed in the process of hammering out a piece of metal which contained numerous segregations consisting of non-metallic impurities or of crystals of some particular alloy phase, segregations such as often occur in a regular pattern in ancient cast bronze of coarse dendritic structure. That bronze of this sort was taken for the fashioning of this strigil handle seems likely. The occurrence of two systems of corroded layers crossing each other may have arisen from the squeezing out of segregated material in two directions by hammering the two pairs of faces of a bar. The squeezing out of segregated non-metallic impurities into layers by hammering seems rather less likely to have occurred than the squeezing out of a particular segregated metal phase into layers by this mechanical means because of the greater plasticity of the segregated metal, especially at high temperatures. That such a segregated metal phase was present in the original piece of bronze from which this handle was fashioned is almost certain because of the high tin content of the alloy. In the cast condition an alloy of this high tin content almost invariably contains considerable delta bronze distributed through the alpha bronze solid solution which constitutes most of the alloy. Since these two different phases have widely different mechanical properties at all temperatures it seems rather likely that they could have been squeezed out into separate alternate layers by the hammering process to which the bronze object was subjected in its manufacture. It is even more likely that the segregated delta bronze was present in the cast alloy in the form of a network between the dendrites of the other phase, and that the processes of annealing and hammering to which the object was subjected merely resulted in the thinning out and the spreading of this network into these layers. Regardless of whether the heterogeneity of the bronze was due to the presence of alternate layers of metal of different composition, to the presence of impurities distributed in layers, or to a combination of both, the two sets of different layers would be acted upon in an entirely different way by water acting as an agent of corrosion, which under these conditions would take place by electrolytic action. The less noble layers in an electrochemical sense would be attacked first, whereas the more noble layers would not then be attacked. Corrosion of the more noble layers would not even start until all the other layers in a given local region had been almost or completely corroded. The appearance of the microstructure at different stages of corrosion appears to show that the two sets of layers were corroded one after the other in just this way.

The views shown in Pl. IV are of the metal in two very advanced stages of general corrosion at the edge of the core. In the top figure considerable remnants of the more resistant cross layers of the metal are still to be seen. In the lower figure, which is a view of a place where corrosion has proceeded further, these remnants of the cross layers have largely disappeared. In all such places at the edge of the core where the corrosion of the metal has become general, the microstructure is very confused. Though parts of grains of metal are still present, and the completely corroded remains of more grains may be readily discerned, much of the original grain structure has disappeared. In the places where the grain structure has disappeared, very irregular and heterogeneous masses of corrosion products are present. As seen under the microscope the color of the intergranular corrosion products in the less corroded parts of the metallic core was almost invariably dark, often brownish or reddish, an indication that cuprous oxide was the predominant compound in these products, but the color of some of the corrosion products in the more corroded parts, such as those pictured in Pl. IV, was light, often whitish or greyish, an indication that stannic oxide was the predominant compound. It seems likely, too, from the color of some of these products, that basic cupric chloride was also present. However, cuprous oxide was still the most abundant single corrosion product in these more corroded parts of the core.

The top figure of Pl. V is a view in cross section at low magnification of one side of the metallic core, most of the patina on this side, and of a transition layer between the core and the patina. In the photomicrograph the metallic core is at the left and the patina at the right. The transition layer shown in this photomicrograph was not present everywhere between the core and the patina proper. It appeared to consist of a dense layer of corrosion products which had no discernible structure, and it was so soft that it could be extruded by the application of moderate pressure when the specimen was clamped in a vise. When the surface of the polished specimen was exposed for some time to moist air a thin line of characteristic bright green basic cupric chloride formed along the line of this transition layer, a phenomenon which was taken as an indication of the presence of cuprous chloride among the corrosion products in this layer. The presence of a considerable proportion of cuprous chloride in these corrosion products would provide a likely explanation for their observed softness.

The color and structure of the patina proper as seen in cross section under the microscope varied much according to the light by which it was viewed and distance from the metallic core. In general, the patina in diffused light had the dark purplish-red color characteristic of cuprite, but the outermost layers, and those next to the metallic core or the transition layer, were much darker in color than the intermediate layers. Some of the intermediate layers were of a bright red color. All these differences in color are probably due to differences in the physical state of the cuprite, especially its degree of subdivision, though they may be due also to differences in the degree to which the cuprite was contamined with other corrosion products. Many specks of white and greenish-blue corrosion products were seen to be distributed throughout the generally dark mass of the patina. That the white specks were composed of stannic oxide seems likely, and it also seems likely that the greenish-blue specks were composed of basic cupric chloride or of a mixture containing a considerable proportion of basic cupric chloride. In its general structure the patina was made up of a large number of successive irregular layers, and the general direction of these was everywhere parallel to the surface of the object. Some of these layers were continuous, but a large proportion were not, and they varied considerably in thickness. This stratified structure is much better shown in the lower figure of Pl. V than in the upper figure. What seem to be layers of light-colored corrosion products in this lower photomicrograph are, for the most part, really empty places between the layers of cuprite or the layers composed largely of cuprite. Such gaps between the layers of the patina were not infrequent. It is important to observe that these layers of corrosion products in the patina lie in a different direction from the layers of corroded and uncorroded metal in the core. By reason of this fact no relationship appears to exist between the layer-like structure in the patina and the layer-like structure in the partly corroded metal.

Apparently the only other corroded Corinthian bronze object ever examined microscopically was a coin of the First Century A.D. This object also was examined by Mr. C. R. Stevens, who studied the microstructure of the corroded parts of the bronze in some detail. Two views of this microstructure have already been published and briefly described.[6] The microstructure of the corroded metal of this coin was similar in some ways to the microstructure of the corroded metal of the strigil handle. Considerable intergranular corrosion also had taken place in the metal of the coin with subsequent progressive corrosion of the grains of the metal. On the other hand, some direct corrosion of the grains without previous intergranular corrosion, an effect not observed in the strigil handle, had evidently occurred in the metal of the coin. Though considerable banded structure was observed in the corrosion products of the metal of the coin, this structure was much less regular and well defined than in the corrosion products of the metal of the strigil handle. Cuprous oxide, much of it occurring in layers, appeared to be also the principal compound present in the corrosion products of the coin, but these corrosion products appeared to contain a larger proportion of other compounds than the corrosion products of the handle. Undoubtedly some of these differences in the microstructures of the corroded parts of these two objects arose from the considerable difference in the chemical composition of the alloys from which they were made.

---

[6] Caley, "The Composition of Ancient Greek Bronze Coins," Amer. Philos Soc., Mem. XI, Philadelphia, 1939, p. 169 and Pl. III.

An explanation of the fact that the early bronze coins of Corinth are usually in a poorer state of preservation than the other Greek bronze coins often found with them at Corinth was advanced a few years ago by D. P. Smith on the basis of his examination of the microstructure of one of these Corinthian coins and of a Sicyonian coin taken as representative of these better preserved coins from elsewhere.[7] The Corinthian coin had a coarse dendritic structure, whereas the Sicyonian coin had a fine worked structure, or, in other words, the metal of the Corinthian coin was very heterogeneous, whereas the metal of the Sicyonian coin was relatively homogeneous. The very heterogeneous structure of the metal of the Corinthian coin was rightly taken as the chief reason why such coins are found in a more highly corroded state than the other bronze coins buried with them and subjected to the same external conditions. Because of the great number of points of different electrical potential on the surface of metal of such marked heterogeneity a great many simultaneous electrochemical reactions occur on this surface on contact of such metal with moist ground or subterranean water, reactions which are destructive to the metal. On the other hand, with metal of greater homogeneity there are fewer points of different electrical potential on the surface, and therefore correspondingly fewer electrochemical reactions which are destructive to the metal. This explanation is undeniably sound, but it is not an explanation as to why bronze objects of all sorts are usually found in a poorer state of preservation at Corinth than at many other Greek sites. It is more than likely that most of these badly corroded objects, particularly the tools and weapons, were composed of metal much more homogeneous than the bronze coins of this city state. As was shown elsewhere[8] from the miscroscopic examination of a series of specimens, the metal of the early bronze coins of Corinth is peculiarly heterogeneous even for Greek bronze coins since these Corinthian coins were evidently struck on cast blanks, a practice that does not appear to have been generally followed at other places.

---

[7] Reported by T. L. Shear, *Am. J. Arch.*, 35, 141–146 (1931).

[8] Caley, *op. cit.*, pp. 165–169.

The abnormal degree to which these coins became corroded is, in part then, the result of a special condition, and the explanation that has been advanced for their unusually poor state of preservation is therefore a special explanation which applies to these coins alone, not to the other kinds of highly corroded bronze objects found at Corinth.

A General Explanation of the Poor State of Preservation of the Bronze Found at Corinth

Any adequate explanation of the extensive corrosion of bronze at Corinth must almost necessarily be based on the general electrolytic theory of metallic corrosion, the theory now commonly accepted as affording a true view of the mechanism by which metals and their alloys become corroded. According to this theory, the active corrosion of a metal or alloy occurs when the surface of the metal or alloy is sufficiently heterogeneous, and when this surface is in contact with a suitable electrolyte. Moreover, such a heterogeneous metal surface in contact with such an electrolyte is considered to have upon it a large number of points or areas which are at different electrical potentials so that numerous local electric currents pass between the members of pairs of these points or areas, and electrochemical reactions of various kinds are produced at or near these places of different electrical potential by the current. Some of these electrochemical reactions, those which occur at the points or areas which are anodic, are destructive to the metal, and various secondary chemical reactions may occur between the products of the primary reactions and the constituents of the electrolyte. The more heterogeneous the metal the greater will be the number of local currents, and the greater will be the rate of corrosion if other conditions remain constant. Likewise, the more heterogeneous the metal in an electrical sense, the greater will be the intensity of these currents, and the greater will be the rate of corrosion.

Heterogeneity in the electrolyte, particularly when it arises from differences in the concentration of dissolved oxygen, may also cause electrolytic corrosion even when the metal itself is very homogeneous. Such heterogeneity may exist in effect when some parts of the metal surface are shielded by the presence of deposits of insoluble corrosion products from the action of the dissolved oxygen while other parts are exposed to the action of this agent. With an ancient alloy such as Corinthian bronze it seems very likely, however, that the marked heterogeneity of the metal itself was a much more important factor in its corrosion than any possible heterogeneity in the electrolyte.

Ancient bronze is usually very heterogeneous, and hence easily corroded, because of the presence in it of many kinds of impurities in considerable proportion. Other conditions being equal, the greater the variety of such impurities, the more heterogeneous the bronze, and the greater the tendency for it to become corroded. From the preceding analyses of specimens of corroded Corinthian bronze it apparently cannot be concluded, however, that such bronze was, by reason of the presence of an unusual variety or an unusually large proportion of impurities, any more heterogeneous than other ancient bronze. Indeed, the actual analytical results indicate that fewer kinds of impurities and lower proportions of these impurities were present than is usual in ancient bronze. Therefore the greater tendency of Corinthian bronze to become corroded cannot be ascribed to any abnormal degree of heterogeneity arising from the presence of impurities. However, the unusually high proportions of tin found in the specimens that were analyzed does seem to indicate that Corinthian bronze was generally more heterogeneous than most ancient bronze for an entirely different reason. In such high-tin bronze a separate alloy phase much richer in tin than the mass of the metal is usually present, this alloy phase, the so-called delta bronze, being ordinarily distributed throughout the body of the metal in the form of numerous minute crystallites or grains. Not only would the existence of such a fundamental heterogeneity in the bronze cause the passage of more numerous electric currents than usual on contact of the surface of the bronze with an electrolyte, but it would cause larger currents than usual to pass because of the relatively large potential differences between the segregations of delta bronze and the mass of alpha bronze constituting the bulk of the alloy. Thus both the greater number and the greater intensity of these local currents would lead to more rapid corrosion than usual. The evidence available appears to show, therefore, that a particular kind of heterogeneity existed in Corinthian bronze, and that this heterogeneity accounts, in part at least, for the greater tendency of this bronze to become corroded.

The nature of the electrolyte which comes into contact with the surface of the bronze also influenzes the rate at which the alloy undergoes corrosion. With buried ancient bronze this electrolyte is usually the local subterranean water. Other conditions being equal, the greater the concentration of dissolved salts in this water the higher the rate of corrosion. Moreover, certain dissolved salts, or rather their ions, will accelerate the corrosion of bronze much more than others. Numerous experiments have shown that chloride ion in particular is effective in furthering the corrosion of copper and its alloys by aqueous solutions.[9] Now, as was shown by an analysis, the present subterranean water at Corinth is unusual in that it contains not only a high concentration of dissolved salts but also chloride ion in considerable concentration. Hence the kind of water which in all probability acted as the electrolyte in the corrosion of the bronze at Corinth was of such a composition that the corrosion of the bronze took place there at a more rapid rate than similar bronze at other sites where the water that acted as the electrolyte had a more normal composition.

The nature of the chemical changes which were involved in the corrosion of the bronze at Corinth was probably to a large extent determined by the chloride ion present in considerable concentration in the subterranean water, the other ions such as sodium ion, calcium ion, and bicarbonate ion being probably of little importance in this respect. The importance of chloride ion in determining the course of the chemical changes in the corrosion of copper and its alloys by solutions containing various dissolved salts has in fact been shown by numerous experimental and theoretical studies, considerable attention having been given to this matter because of the economic importance of the corrosion of such metals by sea water. Though such investigations of corrosion by solutions containing dissolved chlorides have been made almost entirely on pure copper and on alloys of copper and zinc, not on bronzes, and with sea water or with prepared solutions containing rather high concentrations of chloride ion, not on solutions as dilute in chloride ion as the subterranean water of Corinth, the results of these investigations are nevertheless helpful in explaining the nature of the chemical changes which occurred during the corrosion of bronze at Corinth.[10]

---

[9] See especially Gibbs, Smith and Bengough, *J. Inst. Metals,* **15**, 37–191 (1916) and Bengough and Hudson, *J. Inst. Metals,* **21**, 119–166 (1919).

The importance of dissolved chlorides in determining the course of the chemical changes which occur in the corrosion of ancient bronze by brackish waters appears to have first been noticed by Berthelot,[11] who postulated a series of chemical equations for these changes that accounted for the formation of the known copper corrosion products. Berthelot's explanation of the course of the chemical reactions is, however, mostly of historical interest since it is not based on electrochemical considerations, and is in other respects incomplete and incorrect. Also unsatisfactory for the same reasons are the theoretical explanations given by Rosenberg[12] for the nature and course of the corrosion reactions. Much more correct in viewpoint but incomplete and unsystematic are the explanatory remarks of Collins[13] contained in his paper on the corrosion of early Chinese bronzes. A reasonably satisfactory explanation of the course of the chemical changes in the corrosion of ancient buried copper by water or soil containing chlorides is given by Gettens[14] in his paper on the mineralization, electrolytic treatment, and radiographic examination of copper and bronze objects from Nuzi in Iraq. Though much that he writes is applicable also to the course of the chemical reactions in the chloride corrosion of ancient buried bronze, his explanation is apparently intended to apply only to pure copper, not to the alloy bronze. Neither in this explanation nor in any previous one is consideration given to the probable chemical behavior of the tin during the corrosion of bronze by the action of chlorides. The explanation which follows appears therefore to be the first reasonably adequate explanation of the nature and course of the chemical changes that occur when bronze undergoes corrosion for long periods as the result of the action of chlorides contained in soil or subterranean water. Though written primarily as a specific explanation of the chemical mechanism of the corrosion of the bronze buried at Corinth, it is also in a large measure a general explanation or theory of the chemical mechanism of the corrosion of any bronze which has undergone extensive alteration as the result of the action of soil or water containing chlorides.

---

[10] Especially helpful are the experimental studies and theoretical discussion of Bengough and May, *J. Inst. Metals*, 32, 81–142 (1924).

[11] *Ann. chim. phys.* [7] 22, 457–460 (1901).

[12] Antiquités en fer et en bronze, leur transformation dans la terre contenant de l'acide carbonique et des chlorures, et leur conservation, Copenhagen, 1917, pp. 71–80.

[13] *J. Inst. Metals*, 45, 23–55 (1931).

[14] *Technical Studies in the Field of the Fine Arts*, 1, 119–133 (1933).

Important for an understanding of the initial stages in the corrosion of ancient bronze by any natural corrosive agent is the detailed study made by Fink and Polushkin[15] of the microscopic appearance of many specimens of ancient corroded bronze. These investigators found that the places where corrosion started and the paths subsequently taken by the corrosive action varied greatly in accordance with the composition and structure of the alloy, and probably in accordance with the nature of the corrosive agent. In general, however, regardless of where corrosion started, what path it subsequently followed, or the nature of the corrosive agent, the tin in the alloy, not the copper, was always the first to be attacked. This is what would be expected from general electrochemical principles since tin is the more electropositve metal of the two. It was not found, however, that the tin was all corroded away even locally before the corrosion of the copper began, nor even that the parts of the bronze richest in tin were always the first to be attacked. Apparently the potentials of some solid solutions of copper and tin toward electrolytes do not vary directly with the proportions of component metals. For example, in a specimen of cast bronze containing a high proportion of tin, Fink and Polushkin observed that the copper-rich alpha solid solution had been corroded first and that the eutectoid, the component containing the higher proportion of tin, was the last to be corroded. Probably the intermetallic compound of copper and tin contained in the eutectoid caused the grains composed of this solid solution to be more electronegative in character than the grains composed of the alpha solid solution so that the latter were anodic and were therefore corroded by the action of the electrolyte. However, the results of Fink and Polushkin do indicate that in ancient wrought bronze the places richest in tin were usually the first to be attacked. In general these places were the outside layers of the somewhat inhomogeneous grains of alpha solid solution of which such bronzes are composed. Where such places are exposed on the surface of the metal object as edges of grains they are, because of their higher tin content, anodic in character, whereas the central parts of the exposed grains, which are poorer in tin, are cathodic. These edges of the grains are therefore corroded by the action of the electrolyte, whereas the central parts are not corroded. The microscopic examination of specimens of ancient corroded bronze has shown that the corrosion of the outside layers of the grains of metal may extend far below the surface of the object while the central parts of the grains on the surface remain entirely uncorroded. Such, for example, was what had apparently happened initially in the corrosion of the Corinthian strigil handle. From the standpoint of the electrolytic theory of corrosion it is easy to understand why this corrosion of the outside layers of grains may occur far below the surface while most of the surface metal remains uncorroded. The outside layers of the grains are all of the same composition on the average so that the electrical potentials between these layers and the central parts of the surface grains are about the same whether these layers are on grains below the surface or on grains at the surface. Thus corrosion of the outside layers of grains below the surface proceeds in the same way as at the surface providing the electrolyte can penetrate freely down between these grains. Since this so-called intergranular corrosion starts at the surface and proceeds inward with the formation of loose and porous corrosion products and with the formation of open spaces caused by the dissolution of some of these corrosion products, the electrolyte may penetrate more or less easily between the grains as their outside layers are corroded. However, the deeper the electrolyte penetrates into the body of the metal the less easily will it circulate and the longer will spent electrolyte remain in contact with the surfaces of the grains. Thus intergranular corrosion probably tends to proceed at a slower and slower rate as it progresses into the body of the metal. The impurities in ancient bronze, which are often concentrated largely between the grains of metal, probably play an important part in furthering this sort of initial corrosion. Some of these impurities may be cathodic toward the outside layers of the grains of metal with which they are in contact and so tend to accelerate the electrolytic corrosion of these layers. Other impurities may be easily oxidized to soluble corrosion products which are readily dissolved out, and in this way cavities or channels are formed between the grains that allow the electrolyte to penetrate the metal more freely.

---

[15] *Trans. Am. Inst. Min. Met. Eng.*, **122**, 90–117 (1936).

As the anodic parts of the grains of a Corinthian bronze were first attacked by the local water acting as a weak corrosive agent it is likely that the tin was more or less completely dissolved out superficially and that the copper was left behind as a porous layer or as a loose powder. When brass is attacked by weak corrosive agents, the zinc is so removed and the copper is left behind.[16] By analogy a similar process might be expected to take place in the corrosion of bronze by similar agents, though unquestionable experimental proof for the occurrence of this process in the corrosion of bronze, such as has been obtained for brass, does not appear to be available. However, Fink and Polushkin [17] conclude from their observations of the microstructure of ancient corroded bronzes that metallic copper is certainly produced in the initial stages of the corrosion of such alloys. In the corrosion of brass the removal of the more active metal is a well-recognized phenomenon known as *dezincification*, and the analogous process in the corrosion of bronze may be termed *destannification*.[18]

It is highly probable that the porous or powdery copper produced by the removal of the tin from the anodic parts of the grains of the bronze was itself then corroded rather rapidly. As was found by some experiments, filings of even very pure copper are corroded at an appreciable rate by dilute sodium chloride solutions. The more finely divided and impure copper produced by the removal of tin from grains of ancient bronze would be corroded at a much greater rate since the speed of such corrosive action must increase with both the degree of subdivision of the metal and with the degree of its heterogeneity. The almost invariable occurrence of cuprous oxide as an intergranular corrosion product in specimens that are otherwise sound shows that in ancient bronze generally, the corrosion of the copper must begin shortly after the initial attack on the tin. Moreover, the results that were obtained in the experimental corrosion of a specimen of Corinthian bronze show how quickly copper corrosion products appear after the beginning of corrosive attack. Therefore it may be safely concluded that in the natural corrosion of Corinthian bronze, the corrosion of tin and copper began to occur simultaneously soon after the alloy was initially attacked. However, since the tin began to dissolve first, it is logical to discuss first the probable chemical mechanism of the corrosion of the tin in the buried bronze of Corinth by the action of the subterranean water containing chloride ion.

---

[16] Stillwell and Turnipseed, *Ind. Eng. Chem.*, 26, 740–743 (1934).

[17] *Loc. cit.*

[18] Probably to be preferred to the term *detinification* since the verb *detin* in its various forms is already in established use as referring to the removal of tin from scrap tin-plate.

As the tin spontaneously dissolved in the water at anodic areas, the metal must have entered into solution almost entirely as stannous ions, and only to a very slight extent as stannic ions, since it is known from numerous experimental observations that when tin is dissolved by dilute aqueous solutions of all kinds stannous ions are invariably present in the resulting solutions in overwhelming proportion as compared to stannic ions. This process of solution at the anodic areas involved the release of two electrons for each stannous ion that entered the solution. Two electrons then became available at the cathodic areas in the centers of the surface grains, or at other cathodic areas, to combine with the hydrogen ions, or, more strictly speaking, with the hydrated hydrogen ions contained in the water in contact with these areas. The combination may be represented by the equation:

$$2e + 2H^+ \rightarrow 2H.$$

That the hydrogen which was so formed ever appeared as molecular hydrogen gas at these cathodic areas of the metal grains is very unlikely in view of the almost certain presence in the subterranean water of dissolved oxygen which would immediately oxidize the hydrogen to water as soon as it was formed. Indeed, if the hydrogen had not been so oxidized it would have accumulated as a gas on the cathodic areas and probably would have polarized them so that the electrolytic corrosive action could not have proceeded further. That the corrosion of Corinthian bronze soon ceases in a dilute chloride solution free from dissolved oxygen was clearly demonstrated by the last experiment of the three described in a previous section of this paper. The removal of hydrogen ions at the cathodic areas resulted in an excess of hydroxyl ions being left in the solution at these places, two hydroxyl ions for each stannous ion that entered the solution. The final result of the first stage in the corrosion of the tin was, therefore, that stannous ions entered the water at the anodic areas of the grains of metal, and that a corresponding excess of hydroxyl ions accumulated at the cathodic areas of these grains.

As soon as the stannous ions diffused away from the metal surface, or were carried away from it by movement of the water, these ions must have been oxidized rather rapidly to stannic ions by the oxygen dissolved in the water since the instability of stannous ions under these conditions is a well-established fact. This oxidation may be represented by the equation:

$$2Sn^{++} + O_2 + 2H_2O \rightarrow 2Sn^{++++} + 4OH^-.$$

By this oxidation of the stannous ions more electrons were released, two for each ion oxidized, so that still more hydroxyl ions accumulated in the solution, two hydroxyl ions being formed for each stannous ion that was oxidized. In the presence of the chloride ions in the solution, the stannic ions produced by the oxidation probably had little more than a transitory existence for it seems reasonably certain that they combined very rapidly with these chloride ions to form the covalent compound, stannic chloride, a combination that may be represented by the equation:

$$Sn^{++++} + 4Cl^- \rightarrow SnCl_4.$$

Some stannic chloride may have been formed directly by the union of chloride ions with the stannic ions produced in very low concentration in the course of the initial solution of the tin. This direct action probably did not take place to any extent, since if it had done so the equilibria between the stannous ions, the stannic ions, and the metallic tins at the anodic areas would have been shifted to such an extent that few stannous ions would have been left in solution, which appears very unlikely in view of the fact that stannous compounds are present in considerable proportion among the corrosion products of pure tin and alloys of tin which have undergone corrosion from the action of natural waters containing chlorides. The stannic chloride formed in the solution by either of these two ways could have existed but a relatively short time since this compound is readily hydrolyzed by water. This hydrolysis occurs in a series of stages, but the final products are hydrated stannic oxide and hydrochloric acid. The total change that occurred may, therefore, be represented by the equation:

$$SnCl_4 + 4H_2O \rightarrow Sn(OH)_4 + 4H^+ + 4Cl^-.$$

Thus hydrochloric acid was released in the solution, and not only were the original chloride ions restored to the solution for further action but an equivalent number of hydrogen ions were released at the same time, and these may have increased the rate at which the alloy corroded. However, any such increase in rate of corrosion because of increased acidity must have been, at the most, only local and temporary since most of this excess of hydrogen ions was probably soon neutralized by the equivalent excess of hydroxyl ions also present in the solution, either through the diffusion of these ions toward each other or through the movement and consequent mixing of the liquid. Thus the original neutrality of the water in contact with the corroding alloy surface tended to be continually restored in spite of chemical reactions that produced local excesses of hydrogen and hydroxyl ions.

An important effect resulting from the local excess of hydroxyl ions produced at the cathodic areas was very probably the formation of hydrated stannous oxide. By mutual diffusion or through mixing of the solution, the hydroxyl ions produced at these areas came in contact with the stannous ions produced at the nearby anodic areas before these metal ions became oxidized, with the result that insoluble hydrated stannous oxide was formed in formed in accordance with the equation:

$$Sn^{++} + 2OH^- \rightarrow Sn(OH)_2.$$

That hydrated stannous oxide is actually one of the products formed in the natural corrosion of tin has been definitely established by investigations of the nature of the products formed on tin objects which have become corroded by the action of water or moist earth. Moreover, this compound is not a mere transitory product but may persist for long periods as has been shown by analyses of corrosion products found on ancient objects composed of tin, and of alloys of tin. The first analysis of this nature appears to have been made by Bannister,[19] who investigated the nature of the corrosion products on a fragment of a very pure tin scabbard of Medieval, or possibly Roman, date found in England. According to his analysis these corrosion products were composed of 54.68 per cent anhydrous stannic oxide and of 43.35 per cent hydrated stannous oxide, in addition to very small proportions of various compounds other than those of tin. Just recently, Smythe[20] has shown by a series of analyses of corrosion products on ancient tin objects, and on ancient objects of lead-tin alloys, found in Great Britain, that stannous oxide, more or less hydrated, is of general occurrence in these products of corrosion. For example, the layer of corrosion products on a piece of Roman lead-tin alloy, composed of 97.70 per cent tin and 2.73 per cent lead, was found to contain 65.23 per cent stannic oxide, 23.69 per cent stannous oxide, and 3.37 per cent water. An even greater proportion of stannous oxide was found in the corrosion products of a Roman ingot composed of 94.78 per cent tin and 5.37 per cent lead. These contained 56.83 per cent stannic oxide, 32.40 per cent stannous oxide, and 4.25 per cent water. On the other hand, only a very small proportion of stannous oxide was found in the corrosion products of a Roman alloy composed of 97.80 per cent lead and 2.00 per cent tin. In these the stannic oxide amounted to 10.63 per cent and the stannous oxide to only 0.06 per cent. It is interesting to note that chlorine in the form of metallic chlorides was found in most of these corrosion products, an indication that chloride ion was probably instrumental in the corrosion of the tin in these objects just as it was in the corrosion of the tin in Corinthian bronzes. The composition of the corrosion products on an ancient tin ingot dredged up from the sea bottom provides an unquestionable example of the formation of hydrated stannous oxide when tin is corroded in water containing chlorides. These corrosion products contained 84.93 per cent stannic oxide, 5.31 per cent stannous oxide, and 7.00 per cent water. They also contained 1.58 per cent chlorine in the form of metallic chlorides. It seems certain, therefore, that hydrated stannous oxide was formed in the corrosion of Corinthian bronze, and it is very probable that the formation of this compound came about by the mechanism that has been suggested.

---

[19] *J. Inst. Metals,* **35**, 71–72 (1926).

[20] *J. Inst. Metals,* **66**, 355–360 (1940).

The immediate end-product of the corrosion of the tin in Corinthian bronze was, therefore, a mixture of hydrated stannic oxide and hydrated stannous oxide. Smythe[21] has suggested that an actual compound, a hydrated stannous stannate, may be formed between the hydrated higher and lower oxides in the corrosion products of tin, but there is no real evidence that the formation of such a compound takes place under natural conditions. In accordance with its known behavior the hydrated stannic oxide in the mixture lost water gradually. During wet periods this process of dehydration probably never proceeded very far toward completion, but during dry periods it probably proceeded nearly to completion. Under especially favorable conditions and after a long lapse of time the dehydration may have been complete. That the dehydration was ultimately complete, or nearly so, is clearly indicated by some of the analytical results obtained by Smythe in his investigation of the composition of the corrosion products of ancient tin. The hydrated stannous oxide in the mixture also became dehydrated, at least the part of it that was not oxidized beforehand. This dehydration of the hydrated stannous oxide was more rapid and probably more complete than that of the hydrated stannic oxide, since, according to experiments of Bury and Partington,[22] the formation of anhydrous stannous oxide from the hydrate proceeds spontaneously at an easily observable rate even under water at ordinary temperatures. Some of the hydrated stannous oxide was probably oxidized to hydrated stannic oxide by the action of the dissolved oxygen in the subterranean water before its dehydration could occur. As long as the hydrated stannous oxide remained in contact with the metallic tin on the surface of the bronze this oxidation probably proceeded very slowly, but as soon as the compound became separated from the surface by the formation of underlying corrosion products, or as soon as all the tin in a given area on the surface had been corroded, then the oxidation of this compound probably proceeded more rapidly. Likewise, any anhydrous stannous oxide formed by the spontaneous dehydration of the hydrate was probably oxidized sooner or later, the product of this oxidation being either hydrated stannic oxide or anhydrous stannic oxide according to conditions. It seems probable also that stannous oxide, either hydrated or anhydrous, would be less likely to persist among the corrosion products of an alloy containing tin as a lesser component than in the corrosion products of pure tin itself. This appears to be shown by the occurrence of only 0.06 per cent stannous oxide as compared to 10.63 per cent stannic oxide in the corrosion products, analyzed by Smythe, that were found on the ancient lead-tin alloy which contained only 2.00 per cent tin. There is, therefore, little reason to expect much hydrated or anhydrous stannous oxide to be present in the remains of very ancient bronze objects that have undergone very extensive or complete corrosion.[23] The ultimate product, then, of the corrosion of the tin in Corinthian bronze was nearly anhydrous stannic oxide, anhydrous stannic oxide, or a mixture of the two.

---

[21] *Loc. cit.*

[22] *J. Chem. Soc.*, **121**, 1998–2004 (1922).

This mechanism for the corrosion of the tin in Corinthian bronze serves to explain why the ratio of the proportion of tin to the proportion of copper was found to be much lower in the patina of the strigil handle than in the metallic core of this object. Since, by reason of their known behavior, some of the intermediate products of the corrosion of the tin must have entered the water in contact with corroding bronze in either true or colloidal solution, and because of their solubility a considerable proportion of these products were lost by diffusion, or were lost by being carried away mechanically by movement of the water. This must have been particularly true of the stannic chloride and of the products of its hydrolysis, including even the final product, hydrated stannic oxide. Thus a considerable proportion of the tin that was corroded did not appear in the insoluble corrosion products attached to the metallic core. This behavior of the tin is in marked contrast to the behavior of the copper since when the copper was corroded it apparently formed at once a sparingly soluble compound, and though this primary corrosion product was also hydrolyzed, the products of the hydrolysis were compounds still more insoluble than this primary product. Thus very little copper was lost in the course of the corrosion of the bronze. The probable mechanism of the corrosion of the copper in the buried bronze of Corinth by the action of the subterranean water will now be discussed in detail.

---

[23] At the time the analyses were made of the corroded Corinthian bronze objects studied in this investigation, the possible presence of hydrated or anhydrous stannous oxide in their corrosion products was not suspected. Even if the presence of these compounds had been suspected, it is doubtful whether their presence could have been established with certainty, and still more doubtful whether their proportion could have been determined with accuracy, since reliable methods for the detection, and especially the determination, of hydrated or anhydrous stannous oxide in complex mixtures of bronze corrosion products are yet to be devised.

From what has been observed about the behavior of copper when it dissolves in dilute aqueous solutions of various sorts, and from what has been established about the energy relationships involved in such processes, this metal, in contrast to the tin, must have entered into solution during corrosion largely in the higher valence state rather than the lower. According to Allmand and Ellingham[24] the equilibrium constant for the reaction, $Cu^{++} + Cu \rightleftharpoons 2Cu^{+}$, is 0.5×10^{-4} at room temperature, which indicates that the concentration of cupric ions in a dilute copper solution is approximately a hundred times greater than the concentration of cuprous ions when these ions are in equilibrium with a surface of metallic copper. Hence the metal that dissolved at anodic areas in the corrosion of the copper in Corinthian bronze must have entered the water largely as cupric ions and only to a small extent as cuprous ions. This process of solution involved the release of two electrons for each cupric ion that entered the solution. Two electrons then became available at the cathodic areas of the copper to produce, by the same mechanism as in the corrosion of the tin, two hydroxyl ions for each cupric ion that entered into solution. The result of the first stage in the corrosion of the copper was, therefore, that cupric ions, accompanied by an apparently insignificant proportion of cuprous ions, entered into solution at the anodic areas, and that a corresponding excess of hydroxyl ions accumulated at the cathodic areas of the corroding metal.

Although the copper dissolved largely as cupric ion in the course of the corrosion of Corinthian bronze, cupric compounds did not appear as principal final corrosion products, nor were they apparently formed to any extent even as primary corrosion products. The fundamental cause of this apparent anomaly seems to lie in the very low solubility of cuprous chloride in water as compared to that of cupric chloride, a difference that appears to determine the entire course of the chemical changes that occur when copper is corroded by the action of water containing chloride ion in sufficient concentration. Moreover, in water containing chloride ion in sufficient concentration the solubility of cuprous chloride is so reduced by the common ion effect that this salt tends to separate from the solution as an insoluble precipitate even when the cuprous ion concentration is very low. Because of the sharp difference in the solubility of the two chlorides of copper in very dilute chloride solution, it seems very likely on theoretical grounds that cuprous chloride was formed as a principal solid corrosion product at the anodic areas in spite of the fact that the concentration of the cuprous ions at these areas was always very low as compared to that of the cupric ions. In other words, the equilibria between the cupric ions, the cuprous ions, and the metallic copper were so continuously upset by the removal of cuprous ions to form insoluble cuprous chloride that this salt was formed as a principal corrosion product in spite of the fact that the metallic copper tended to dissolve almost entirely as cupric ions. Various observations and facts support this explanation of how cuprous compounds began to be formed as corrosion products instead of cupric compounds.

---

[24] *The Principles of Applied Electrochemistry*, London, 1924, p. 262.

Though the presence of cuprous chloride was actually observed among the corrosion products of only one of the corroded Corinthian bronze objects, and this not with entire certainty, the fact that cuprous chloride is formed as a primary corrosion product when such bronze is corroded by water containing chloride ion was clearly demonstrated by the results of the experiments on the corrosion of a specimen of this bronze. The formation of cuprous chloride as a primary corrosion product when pure copper, itself, is corroded by the action of dilute chloride solutions was observed by Evans.[25] That the same compound is formed as a primary product in the corrosion of copper by sea water was observed by Bengough and May.[26] A few investigators have definitely identified cuprous chloride as a corrosion product of ancient bronze and copper objects which were known to have been corroded, or which were presumably corroded, in the presence of water containing chlorides. The earliest investigator to notice the presence of cuprous chloride among the corrosion products of ancient bronze objects appears to have been Rosenberg,[27] who, by an examination of its physical properties and by means of analyses made for him by Baggesgaard-Rasmussen, was able to identify as cuprous chloride a distinctive corrosion product isolated from certain ancient bronzes. Gettens [28] appears to have been the first to identify cuprous chloride as a corrosion product of ancient copper objects. In some of the specimens of corroded copper nails about 3500 years old from Nuzi in Iraq which he examined, cuprous chloride was found to be a principal component of the corrosion products. The present writer found cuprous chloride to be a principal component of the corrosion products of many of the extensively corroded bronze and copper objects recovered from deep wells at the site of the Athenian Agora. It seems likely, incidentally, that the poor state of preservation of these objects may be explained in much the same way as that of the bronze objects of Corinth since the water of these wells, like that of the Fountain Peirene, contains chloride ion in much higher concentration than is usual in subterranean water. Most of the cuprous chloride in or on the objects from the wells of the Athenian Agora was soft, colorless, and translucent without visible crystalline form, some of it was soft, translucent, and without crystalline form but was stained various shades of blue or green from the presence of small amounts of admixed cupric compounds, and some of it occurred in the form of small, brilliant, colorless, translucent, tetrahedral crystals, which corresponded in every respect to the crystalline form of the rare mineral, nantokite.[29] That cuprous chloride is a corrosion product of ancient copper, when this metal is corroded in the presence of water containing chloride ion, must, therefore, be considered an established fact. That the same compound was formed in the course of the corrosion of the copper in Corinthian bronze seems certain also.

---

[25] *Metallic Corrosion, Passivity, and Protection,* London, 1937, p. 365.

[26] *J. Inst. Metals,* 32, 101–102 (1924).

It was observed both by Gettens and the present writer that the cuprous chloride in the corrosion products of ancient bronze and copper objects almost invariably occurs next to uncorroded metal and never in the outermost layers of these corrosion products. This, and the fact that cuprous chloride has never been found in the corrosion products of completely corroded bronze and copper objects, show clearly that cuprous chloride is an initial and unstable corrosion product, not a final and stable product of corrosion. Indeed, cuprous chloride is such an unstable compound in the presence of oxygen and water that its survival is not to be expected under normal conditions even in objects that contain much uncorroded metal. In the presence of an abundant supply of oxygen and of moisture, cuprous chloride is so very unstable that its formation is probably only transitory in chloride solutions in which an abundant supply of air reaches the corroding metal. This was apparently what happened in the first experiment on the corrosion of a specimen of Corinthian bronze where, because of frequent exposure of the corroding metal to the air, no visible amount of cuprous chloride was formed. On the other hand, in the second experiment, where the supply of oxygen to the metal was deliberately restricted by keeping the bronze specimen submerged in the chloride solution and by restricting the supply of air to the surface of this solution, cuprous chloride was formed in visible amount. When copper is corroded in porous soil by the intermittent action of subterranean water containing chlorides, or in any other situation where oxygen may freely reach the surface of the corroding metal, cuprous chloride cannot, therefore, be expected to survive long as a corrosion product. Necessary also for the formation of cuprous chloride as more than a mere transitory product is the corrosion of the copper in a somewhat restricted volume of water, since a large volume of water, especially if renewed constantly, would certainly dissolve the cuprous chloride as fast as it was formed in spite of the low solubility of this salt in water. Moreover, cuprous chloride is not only dissolved but is actually decomposed by an abundant excess of water. Hence the formation of cuprous chloride as more than a transitory corrosion product is not to be expected in flowing water, nor in a large volume of stagnant water, even when the supply of oxygen reaching the surface of the corroding metal is meager. Obviously the ideal condition for the survival of cuprous chloride as a corrosion product of copper in water containing chloride ion is the submersion of the metal in a small volume of such water away from free access to air. This very condition appears to account for the occurrence of cuprous chloride in relatively large amount as a corrosion product of the bronze and copper objects recovered from among the debris near the bottom of the deep wells of the Agora. Such an ideal condition for the survival of cuprous chloride as a corrosion product is obviously not of frequent occurrence, and probably did not exist for the bronze buried in the ground at Corinth. Therefore, the general absence of cuprous chloride in the actual corrosion products of Corinthian bronze objects is not surprising.

---

[27] *Op. cit.*, pp. 70–72.

[28] *Loc. cit.*

[29] It seems scarcely necessary to point out that these observations on the formation of cuprous chloride in the corrosion of ancient copper objects show conclusively that the natural cuprous chloride, nantokite, may originate from the action of saline or slightly saline water on native copper.

The second stage in the corrosion of the copper in Corinthian bronze was, then, the formation of cuprous chloride at the anodic areas of the corroding metal in accordance with the equation:

$$Cu^+ + Cl^- \rightarrow CuCl.$$

In spite of the absence of cuprous chloride in the actual corrosion products of Corinthian bronzes, it nevertheless seems probable that nearly all the copper in such bronzes did enter into combination as cuprous chloride during the course of their corrosion. This not only seems probable from the low solubility of this salt and from the probable effect of this low solubility on the ionic equilibria at the anodic areas of the corroding copper, but it also seems probable from the nature of the actual corrosion products and from the mode of occurrence of cuprous chloride that has survived in the corrosion products of ancient bronze and copper objects which have undergone corrosion in water containing chloride ion. For example, in many such objects from the wells of the Agora, the corroding copper was not only completely surrounded by cuprous chloride but no other corrosion products were present in the layers of cuprous chloride next to the corroding metal. This mode of occurrence certainly indicates that the cuprous chloride was formed as the sole initial solid corrosion product and that all the other corrosion products were formed from this salt. In view of the nature of the corrosion products now found on Corinthian bronzes, and in view of the ground conditions at Corinth, it seems probable, moreover, that the cuprous chloride formed during the corrosion of such bronzes was decomposed soon after it was formed.

The decomposition of cuprous chloride in the presence of water and oxygen occurs both by hydrolysis and oxidation. That the rate of decomposition and the nature of the products varies greatly according to the conditions under which the decomposition takes place has been shown experimentally by Gröger,[30] and especially by Bengough and May.[31] In the decomposition of cuprous chloride by distilled water containing oxygen in the form of dissolved air, the nature of the products varies according to temperature, time, concentration of oxygen, and the ratio of the mass of the salt to the mass of the water. Since the decomposition of cuprous chloride in natural corrosion processes takes place at nearly a fixed temperature, and since the time factor may be considered indefinitely long, the concentration of dissolved oxygen and the ratio of the masses of the salt and the water are the important factors from the standpoint of the present investigation. When the water is present in ample excess, only the concentration of the dissolved oxygen is important.

That cuprous chloride is readily decomposed even by water free from oxygen was shown by Gröger, who investigated the nature of the products formed when the salt is treated with successive portions of freshly boiled water in an atmosphere of hydrogen or of carbon dioxide. The residue resulting from the complete decomposition of the salt by such treatment was found to consist mostly of cuprous oxide. The effect of different concentrations of dissolved oxygen was studied by Bengough and May in the following manner. A small fixed amount (0.5 g.) of cuprous chloride was placed in each of a series of vessels of decreasing diameters, and a fixed volume of ordinary distilled water was added to each vessel. In this way the cuprous chloride was immersed in increasing depths of water, and as a consequence atmospheric oxygen was less and less readily available to react with the cuprous chloride or its hydrolytic decomposition products. In the vessel in which the depth of water was the least (2.6 cm.), basic cupric chloride was formed as the sole solid decomposition product. In all the others, cuprous oxide was formed in addition to basic cupric chloride. Bengough and May also studied the effect of different ratios of mass of water to mass of salt, though these experiments necessarily involved to some extent the effect of different concentrations of dissolved oxygen. Volumes of distilled water ranging from 2 to 500 cc. were added to the same fixed amount (0.5 g.) of cuprous chloride, and the composition of the solid residues was determined after 24 hours. In the two experiments in which the smallest volumes of water were used, no cuprous oxide could be detected in the residues, which appeared to contain only basic cupric chloride as a decomposition product. In all the other experiments cuprous oxide was present in the residues, and it was present in increasing proportion as the volume of water increased. In the experiment in which the largest volume of water was used, cuprous oxide was almost the sole solid decomposition product, only a slight amount of basic cupric chloride being formed. The decomposition of the cuprous chloride was incomplete in most of these experiments at the time they were terminated. In another experiment by Bengough and May in which 0.5 g. of cuprous chloride was treated with a large volume (400 cc.) of water, and in which the supply of atmospheric oxygen reaching the surface of the liquid was restricted, the decomposition of the salt was complete at the end of six days, and the final decomposition product was found to consist mainly of cuprous oxide, the remainder being basic cupric chloride. In all these experiments about a third of the copper in the decomposed cuprous chloride was oxidized to soluble cupric ion. Apparently when the supply of oxygen is restricted, the decomposition of the salt proceeds largely by hydrolysis and autoxidation, but when the supply of oxygen is ample, the decomposition proceeds largely by oxidation processes, not by hydrolysis. Though the pronounced effect of large volumes of water in determining the nature of the decomposition products is considered by Bengough and May to be mainly the result of a restriction of the supply of atmospheric oxygen, it seems rather more likely that the effect of a large excess of water is mainly that of furthering the hydrolytic decomposition of the salt by making possible the free diffusion of the hydrogen and chloride ions, formed on hydrolysis, away from the immediate neighborhood of the decomposing salt, so that these products do not accumulate there in considerable concentration and thus hinder the further decomposition of the salt by hydrolytic action.

---

[30] *Z. anorg. Chem.*, 28, 154–161 (1901)

[31] *J. Inst. Metals*, 32, 108–120 (1924).

The conditions maintained in the particular experiments of Bengough and May just described are not however the conditions which actually exist during the decomposition of the cuprous chloride formed in the course of the natural corrosion of copper. In the decomposition of cuprous chloride that is formed as a natural corrosion product the water contains chloride ion and the decomposition occurs in the presence of a surface of metallic copper. Bengough and May showed by further experiments that the chloride ion concentration may be important in determining the nature of the ultimate products of the decomposition of cuprous chloride in water containing dissolved oxygen. When the chloride ion concentration is very high, as in sea water, cuprous oxide does not appear as an ultimate solid decomposition product, the product always being basic cupric chloride. Evidently a very high concentration of chloride ion so represses the hydrolytic decomposition of the cuprous chloride that no cuprous oxide is formed. With decreasing concentrations of chloride ion more and more cuprous oxide and less and less basic cupric chloride appear as decomposition products, provided always that the volume of the water is large. Thus, for example, on treating a 0.5 g. portion of cuprous chloride with a large volume of a one hundredth of one per cent solution of sodium chloride, Bengough and May found that cuprous oxide was formed in large proportion just as in distilled water. Hence with concentrations of chloride ion as low as that in the subterranean water at Corinth the effect of chloride ion is not very important. In the presence of a copper surface, cupric ions do not appear in noticeable amount as a decomposition product of the cuprous chloride, as was also shown experimentally by Bengough and May. In their experiments discs of copper foil were placed on the bottoms of beakers and various amounts of cuprous chloride were placed on these discs, after which 100 cc. of distilled water was added to each beaker, and watch glasses were placed on the beakers. At the end of twenty days the one disc which had been covered with a heap of cuprous chloride had entirely disappeared, and the solution contained cupric ion in noticeable amount. In the two others where less cuprous chloride was present the copper had not corroded away in this time, nor was it corroded away on continuing the experiments for twelve days longer. At the end of this second period of standing, the solution in the vessels was tested for the presence of cupric copper by the addition of ammonium hydroxide but no reaction was observed. The product of the decomposition of the cuprous chloride in these two experiments was a mixture of cuprous oxide and basic cupric chloride. As long as metallic copper is present, these then are the decomposition products of cuprous chloride when this salt is decomposed by the action of an ample excess of water containing dissolved oxygen.

From the results of all these experiments it is to be expected, therefore, that cuprous oxide and basic cupric chloride would be the ultimate products of the decomposition of the cuprous chloride formed in the course of the corrosion of the copper in Corinthian bronzes, and these very products, it will be noted, are the ones which were found by analysis to be present in the corrosion products of such bronzes. Moreover, since the analyses show that cuprous oxide is present in these corrosion products in overwhelming proportion as compared to basic cupric chloride, the results of these experiments indicate that the decomposition of the cuprous chloride probably took place by the action of considerable quantities of water. This in turn seems to indicate that the corrosion of the bronzes occurred in soil often saturated with subterranean or surface water, not merely moistened with it. In general, the relative proportion of cuprous oxide and basic cuprous chloride in the corrosion products of ancient bronze and copper objects which have undergone corrosion by the action of water containing chlorides may provide a valuable clue to the ground conditions that prevailed during the corrosion of such objects.

From the results of the experiments which have been described, and from the results of the analyses which were made of the corrosion products of Corinthian bronze, it seems very probable, then, that the cuprous oxide which occurs as the principal corrosion product was produced from cuprous chloride largely by the hydrolytic decomposition of this salt. This hydrolytic decomposition reaction may be represented by the equation:

$$2CuCl + H_2O \rightarrow Cu_2O + 2H^+ + 2Cl^-.$$

As indicated by the equation, hydrochloric acid was released in the solution by this hydrolysis of the cuprous chloride, and not only were the original chloride ions restored to the solution but an equivalent number of hydrogen ions were released at the same time, and these may have tended to increase, locally and temporarily, the rate of the corrosion of the bronze in the same way as the hydrogen ions released by the hydrolysis of the stannic chloride produced in the course of the corrosion of the tin. On the other hand, as long as these hydrogen ions remained in the immediate neighborhood of the decomposing cuprous chloride, the hydrolysis of this salt probably could not proceed to completion. However, these hydrogen ions probably did not accumulate in the neighborhood of the decomposing salt and so hinder its complete hydrolysis since they probably diffused away or were carried away by the movement of the water, or, more probably still, were soon neutralized by the hydroxyl ions present in equivalent amount at the cathodic areas of the corroding copper, either through the diffusion of these ions toward each other or through the movement and consequent mixing of the liquid.

Though some cupric ions arising from the dissolution of the metal or from the autoxidation of cuprous ions were always present in the film or layer of solution around the corroding copper, additional cupric ions were undoubtedly formed to some extent by the oxidation of cuprous ions by the oxygen dissolved in the subterranean water, especially in parts of the film or layer of solution not immediately next to the metal surface. This oxidation may be represented by the equation:

$$4Cu^{+} + O_2 + 2H_2O \rightarrow 4Cu^{++} + 4OH^{-}.$$

The hydroxyl ions so produced must have united almost at once, however, with part of the cupric ions to form slightly soluble hydrated cupric oxide in accordance with the equation:

$$Cu^{++} + 2OH^{-} \rightarrow Cu(OH)_2.$$

Some hydrated cupric oxide may also have formed by the union of cupric ions with hydroxyl ions produced at the cathodic areas of the corroding copper. In the presence of the other cupric ions and the chloride ions present in abundance in the solution, this compound probably did not long exist by itself but combined with some of these ions to form the more stable and still more insoluble basic cupric chloride, atacamite, a transformation which may be represented by the equation:

$$3Cu(OH)_2 + Cu^{++} + 2Cl^{-} \rightarrow 3Cu(OH)_2 \cdot CuCl_2.$$

That hydrated cupric oxide actually reacts readily with cupric ions and chloride ions to form this basic cupric chloride was demonstrated by Sabatier,[32] who found that he could easily prepare this particular basic salt by treating the hydrated oxide with a cold dilute solution of cupric chloride.

From the frequent occurrence of basic cupric chloride as the principal corrosion product on the exterior of highly corroded bronze and copper objects in which cuprous oxide is the predominant substance in the corrosion products as a whole, it is probable that some basic cupric chloride is often formed by the alteration of cuprous oxide. Probably cuprous ions arising from the dissolution of cuprous oxide in a film or layer of water surrounding a buried object of this sort are oxidized by air or dissolved oxygen with the subsequent formation of the basic chloride by much the same mechanism as was suggested for its more direct formation. Because of the very low solubility of cuprous oxide, this change must, however, be very slow in spite of the unusually favorable conditions for the oxidation of cuprous ions at the outside surface of a highly corroded object. How slowly this transformation really occurs is indicated by the fact that the layer of basic cupric chloride on the outside surface of such objects, such as that found on the Corinthian strigil handle, is often very thin even after centuries of corrosive action.

In previous explanations of the mode of formation of basic cupric chloride as a corrosion product of ancient bronze or copper, it has been generally assumed that molecular cuprous chloride was oxidized directly to basic cupric chloride, and even that cuprous oxide was simultaneously involved in the formation of this basic salt. These explanations have usually included the formulation of complicated equations with large numerical coefficients before the formulas of the reacting substances. Since the correctness of such equations is inherently very improbable, it seems to the writer that the present simpler explanation is considerably nearer the truth.

Because of their insolubility and marked chemical stability, these two corrosion products of Corinthian bronze, cuprous oxide and basic cupric chloride, underwent little or no transformation into other compounds in the course of the centuries, even into the very stable basic copper carbonate, malachite. A small amount of this basic copper carbonate, which apparently had originated from the alteration of one of these products, was indeed detected on the outside surface of the corroded strigil handle, and a trace of carbonate was detected in the corroded remains of the strigil blade, but in general the analytical examination of the corrosion products of the various Corinthian bronze objects yielded little indication of the transformation of the cuprous oxide or the basic cupric chloride into this basic copper carbonate or into any other copper compound. It is interesting to note that Gettens[33] also found but a trace of carbonate in the highly corroded remains of some very ancient copper nails, which, from the presence of cuprous oxide and copper chlorides as principal corrosion products, evidently had been corroded in the presence of water containing chlorides. The results of the present investigation certainly show that basic copper carbonate is not formed as a primary corrosion product in the presence of water containing chloride ion in sufficient concentration, even when much bicarbonate ion is also present in the water, and they show, moreover, that basic copper carbonate is formed only in small amount as a secondary product when copper is corroded under these conditions. This absence of basic copper carbonate as a corrosion product may perhaps be surprising in view of the general opinion among chemists, as indicated by their statements in numerous chemical works, that basic copper carbonate, especially in the form of malachite, is invariably a principal product of the corrosion of copper under natural conditions. The absence of basic copper carbonate as a corrosion product of Corinthian bronze is, however, by no means a special phenomenon. Examination by the writer of the corrosion products of a considerable number and variety of ancient bronze and copper objects has convinced him that this compound, contrary to all such statements, does not occur frequently as a principal component of the corrosion products of ancient bronze and copper objects, and in fact is often not present at all in sensible amount. Moreover, as has been shown by some recent investigations,[34] basic copper carbonate is not even the principal component of the green patina formed on bronze and copper objects which have been exposed for a long time in the open air, the principal component of this green patina being basic copper sulfate, basic copper chloride, or an intimate mixture of the two. The truth appears to be that no sound experimental basis exists for the various statements which have been made as to the invariable formation of basic copper carbonate as the principal product of the corrosion of copper under natural conditions.

---

[32] *Compt. rend.*, 125, 103 (1897).

---

[33] *Loc. cit.*

[34] See especially, Vernon and Whitby, *J. Inst. Metals,* 42, 181–202 (1929), and Vernon, *ibid.,* 52, 93–100 (1933).

The chemistry of the corrosion of the tin and the chemistry of the corrosion of the copper in Corinthian bronze have now been considered separately as though the chemical changes which occurred during the corrosion of the one metal took place quite independently of those which occurred during the corrosion of the other. In reality, however, these two sets of chemical changes did not take place independently, the chemical properties of the two metals and the chemical properties of some of the corrosion products of these metals being such that certain chemical interactions must have occurred. In general, chemical interactions could have occurred between one of the metals and certain of the corrosion products of the other, between solid solutions or compounds of the two metals and certain corrosion products of either metal, and between certain corrosion products of the two metals. As a specific example of such interaction, it is very probable that any cupric ions which came into contact with metallic tin on the surface of the alloy, or with a solid solution or compound of tin and copper which had the proper potential toward such ions, were at once reduced to metallic copper with the resultant production of an equal number of stannous ions in accordance with the equation:

$$Cu^{++} + Sn \rightarrow Sn^{++} + Cu.$$

In this way some copper which had already dissolved was precipitated on the surface of the corroding alloy. That such precipitation of copper does occur is at least strongly indicated by the occasional presence of bright crystals of copper next to, or near to, the metallic core of bronzes which have undergone severe corrosion. Excellent specimens of such crystals were observed by the writer among the corrosion products of deeply corroded bronzes taken from some of the wells of the Athenian

Agora. Of course it must be admitted that the existence of such copper crystals is not conclusive proof that they originated in this particular way since they also could have originated by the auto-reduction of cuprous chloride or by the action of local currents generated from inequalities in the concentration of copper ions in the liquid around the corroding alloy. Sooner or later such precipitated copper must have again dissolved and passed through the same stages in corrosion as any other copper in the alloy, and, in general, the substances formed by any of these chemical interactions between metals and corrosion products or between the corrosion products themselves probably passed ultimately through the same stages in corrosion as any of the other substances. In other words these chemical interactions did not really alter the general course of the chemical changes which occurred during the corrosion of the tin and the copper. For this reason it does not seem necessary to discuss all these possible interactions in detail.

Little consideration also need be given here to the chemical changes which occurred in the course of the corrosion of the minor components of early Corinthian bronze such as the lead and the iron. As shown by the analytical results given in the first part of this paper, these components of the alloy were so small in amount as compared to the tin and the copper that the changes which occurred during their corrosion probably had little effect on the main course of the corrosion of the bronze or on the composition and structure of the principal products of corrosion. Moreover, the difficulty of identifying experimentally either the intermediate or the final products of the corrosion of these minor components makes difficult any explanations of the chemical changes which occurred during their corrosion that are as soundly grounded upon observed facts as are the explanations which are given for the chemical changes that occurred during the corrosion of the tin and the copper. However, the behavior of elementary lead and iron, when acted upon by waters of the most varied composition, has been so thoroughly studied by so many investigators that the changes which occurred during the corrosion of these metals as components of Corinthian bronze may be postulated with some degree of confidence. Thus it is reasonably certain that the lead entered into solution almost entirely as plumbous ions with the production of an equivalent amount of hydroxyl ions. Some of hese lead ions may then have combined with chloride ions in he water to form sparingly soluble lead chloride as an intermediate corrosion product, though, because of the small amount of lead in the alloy, the consequent very small amount of lead in solution at any given time, and the appreciable solubility of lead chloride, it seems rather unlikely that any of this salt was so formed. It is much more likely that all the lead ions combined almost at once with hydroxyl ions to form insoluble hydrated lead oxide as an intermediate corrosion product. By the action of the bicarbonate and the carbonate ions in the water this intermediate product was then transformed more or less rapidly into basic lead carbonate, the final product of the corrosion of the lead. Much of the iron in Corinthian bronze objects may have actually been in the form of ferric oxide before corrosion started, either because it entered the bronze in this form or because it was oxidized to this compound during the preparation of the bronze or during the fabrication of the objects from the prepared alloy. In the course of corrosion any elementary iron in the bronze must have entered into solution largely as ferrous ions with the production of an equivalent quantity of hydroxyl ions. Most of these ferrous ions were probably oxidized rapidly to ferric ions by the dissolved oxygen in the subterranean water. Some of these ions may have combined first with hydroxyl ions to form nearly insoluble hydrated ferrous oxide as an intermediate corrosion product. The ferric ions produced by the direct oxidation of ferrous ions, or those produced by the oxidation of ferrous ions which had first combined to form hydrated ferrous oxide, then combined rapidly with hydroxyl ions to form a hydrated ferric oxide as the immediate final corrosion product. This compound probably lost water gradually until a less hydrous ferric oxide was formed as the final product of the corrosion of the iron. Some chemical interactions involving these minor components and their corrosion products probably took place, but these interactions were probably insignificant as compared to those which took place between the major components of the alloy and their corrosion products.

Though the various chemical reactions which occur during the corrosion of buried bronze constitute the most important changes that take place in the transformation of metal into corrosion products, these reactions are by no means the only changes which occur. Various physical changes must also occur, such as, for example, changes in the size of the individual particles of the corrosion products, changes in the state of aggregation of such particles, and changes in the relative positions of the different kinds of particles and aggregates.

Though the general composition of the corrosion products or patina of an ancient bronze may be satisfactorily explained on the basis of chemical changes, the arrangement of the individual corrosion products in relation to each other, or the general structure of the patina, often cannot be satisfactorily explained on the basis of such changes alone, but must be explained in part on the basis of physical changes. Since the occurrence of a particular physical change, or of a particular series of physical changes, during the corrosion of an ancient bronze cannot ordinarily be established with the same degree of certainty as the occurrence of a particular chemical change, or of a particular series of chemical changes, the existence of a given type of structure in the patina of an ancient bronze is usually not so easy to explain as is its general composition. However, certain types of structure may perhaps be adequately explained on the basis of chemical considerations. For example, the fine stratified structure of the patina of the Corinthian strigil handle, which was described in some detail before, and of which a magnified view is given in the lower figure of Pl. V, is a type of structure that may be so explained. Such a stratified structure may be the result of periodic general corrosion, a type of corrosive action in which all the processes of corrosion proceed freely at a time when conditions are favorable, and then stop when conditions become unfavorable, only to begin again, and stop again, and so continue alternately in accordance with the alternations in the conditions. Thus in the corrosion of buried bronze all the processes of corrosion may proceed apace during periods of wet weather but may cease entirely during periods of dry weather. The very thin layer of fresh corrosion products formed on the metal during a given wet period, when much subterranean water is present, may become dried out when the ground again becomes dry, and, in doing so, shrink and become partly detached from the metal surface and from the previous layer of corrosion products. In this way many very thin layers may be formed in succession which are not only distinct from each other but may be separated in part by minute gaps such as those observed between many of the layers in the patina of the strigil handle. Instead of corrosion being periodic because of the alternate cessation and resumption of all corrosive action it may be periodic because certain chemical reactions proceed more readily under one set of conditions and others more readily under another set of conditions. In other words corrosion proceeds at all times, but, in accordance with the alternations in the conditions, certain reactions will predominate and determine the course of the corrosion and the nature of the final products at one time, and other reactions will predominate and determine the course of the reactions and the nature of the final products at another time. Thus, as was explained before, the decomposition of cuprous chloride during the corrosion of Corinthian bronze probably did not proceed rapidly when the amount of water in contact with the buried bronze was small but did proceed rapidly when the amount of water was large. Furthermore, as was also explained before, basic cupric chloride was probably formed to a larger extent as a final corrosion product and cuprous oxide to a smaller extent when the amount of water was small, and the reverse was probably true when the amount of water was large. In this way, and in general in similar ways, successive thin layers of corrosion products were produced which alternated in composition in accordance with the alternations in the conditions. Moreover, these thin layers probably became more distinct in the course of time as subterranean water preferentially leached out the more soluble corrosion products from the patina. By such action the minute gaps observed between the layers of the patina of the strigil handle could also have been formed. From the nature of the climate it is likely that both these types of periodic corrosive action occurred during the corrosion of the bronze buried at Corinth, and it seems likely that both types of action were of frequent occurrence in the corrosion of ancient bronze buried in many other localities. This explanation of the origin of a fine stratified structure in the patina of ancient bronze seems more likely to be true than a possible alternative explanation of the origin of such a structure by the differentiation of an original more or less homogeneous single layer of corrosion products into many layers of different or similar composition by some sort of periodic precipitation process arising from the dissolution and diffusion of certain of the substances in the layer.

Since the formation of a fine stratified structure in the patina of an ancient bronze appears to be connected with seasonal or climatic changes it seems possible that the number and thickness of the layers in the patina may somehow serve as an index of the length of time the bronze was buried, or of the prevalence of certain climatic conditions at certain times during the period of burial. However, it seems too much to hope that the number and thickness of the layers in the patina of an ancient bronze can ever be of such definite chronological and climatic significance as the number and thickness of the annual rings in the cross section of an ancient log. Too many factors of uncertain or unknown effect are probably involved in the formation of these layers. Hence the possibility of attaching any chronological significance to the number of these layers must exist only for objects of the same composition buried at a given site, and even for these the number of the layers may be *only* of relative and not of absolute chronological significance. In spite of these uncertainties the possibility of attaching some chronological significance to the number of layers, in special circumstances at least, may be well worth some investigation.

AN INTERPRETATION OF THE ALLUSION OF PAUSANIAS TO THE TREATMENT OF BRONZE IN THE WATER OF THE FOUNTAIN PEIRENE [35]

One of the few allusions made by Pausanias to technological matters in his famous *Description of Greece* is contained in the short passage about the Fountain Peirene at Corinth [36] in which he incidentally remarks that the bronze of Corinth was treated by means of the water of this fountain. The text of this particular part of the passage according to Spiro [37] is as follows:

καὶ τὸν Κορίνθιον χαλκὸν διάπυρον καὶ θερμὸν ὄντα ὑπὸ ὕδατος τούτου βάπτεσθαι λέγουσιν, ἐπεὶ χαλκός . . . γε οὐκ ἔστι Κορινθίοις.

---

[35] Discussions with B. D. Meritt and O. Broneer of the Institute for Advanced Study about the meaning of certain Greek words were of help to the author in writing this part of the paper.

[36] Book II, Chapter 3.

[37] *Pausaniae Graeciae Descriptio,* Leipzig, 1903, Vol. I, p. 130.

The difficulties which have been encountered in the interpretation of this interesting statement may perhaps be most quickly and strikingly indicated by the very different wording of the translations of it which appear in the three most recent English versions of this work of Pausanias. The oldest of these translations is that of Shilleto,[38] which reads as follows:

> and they say that Corinthian brass when hissing hot is dipped into this water.

Frazer[39] gives the following translation:

> and they say that the so-called Corinthian bronze gets its color by being plunged red-hot into this water; for, in point of fact, Corinth has no bronze of its own.

The most recent translation, that of Jones,[40] reads as follows:

> and they say that the Corinthian bronze, when red-hot, is tempered by this water, since bronze . . . the Corinthians have not.

The purpose of the present remarks is to show that some of the difficulties, at least, which have been encountered in the interpretation of this statement of Pausanias may be resolved by the aid of certain previously known technological facts and from the results of some experiments made by the writer, especially when these facts and the results of these experiments are considered in connection with the other kinds of evidence which are of help in the determination of the meaning of this statement.

Undoubtedly the principal difficulty in the interpretation of the first phrase of this statement, the part preceding ἐπεί, is the intended meaning of βάπτεσθαι, an infinitive form of βάπτω, the primary meaning of which is *dip*, more especially with reference to the dipping of something in water or some other liquid.[41] This is the meaning taken by Shilleto for his translation, but there are some very good reasons for concluding that this is not the intended meaning in this particular context. In the first place, as has been previously pointed out by Frazer[42] in refutation of a statement of Curtius, the phrase ὑπὸ ὕδατος τούτου itself indicates the dipping or immersion of the bronze in the water, so that something more than the bare sense of dipping is to be attached to βάπτεσθαι, otherwise the statement of Pausanias appears redundant. Moreover, in most of its actual uses the verb βάπτω, especially in the passive voice, implies something more than the mere act of dipping. It implies dipping for some particular purpose, and more often than not the verb takes on a derived or secondary meaning indicative of this purpose. If Pausanias meant to use the verb only in its primary sense it then follows that he is merely mentioning the act of dipping the bronze without giving any hint as to the purpose of this act, which seems rather unlikely. Furthermore, the second phrase of this statement of Pausanias, the part beginning with ἐπεί, which is certainly to be interpreted as meaning that the bronze was not of local origin, seems to imply that the metal was brought to Corinth for working or for treatment with the water of the Fountain Peirene because of some unusual quality in this water. Though it must be admitted that this implication is rather vague or awkwardly phrased in the text as it is now known, this apparent vagueness or the apparent awkwardness of phrasing may be chiefly the result of the lacuna considered by some editors to exist in the text at this point. Shilleto avoids the difficulty inherent in the exact interpretation of this second phrase, and its apparent bearing on the meaning of the first phrase, by the convenient device of ignoring it altogether in his translation. That βάπτεσθαι meant that the bronze was merely dipped or plunged does not, therefore, seem likely, and it was probably for reasons such as those just given that this interpretation of its meaning in this context was rejected by the two translators of Pausanias who followed Shilleto. However, as is shown by their translations, Frazer and Jones attach radically different meanings to βάπτεσθαι both of which are possible from the way the verb βάπτω is used by other writers. The problem as to which of these two meanings is the more likely will now be considered.

---

[38] Pausanias' Description of Greece Translated into English with Notes and Index, London, 1886, Vol. I, p. 95. (Bohn's Library.)

[39] Pausanias' Description of Greece Translated with a Commentary, London, 1898, Vol. I, p. 74.

[40] Pausanias' Description of Greece with an English Translation, London and New York, 1918, Vol. I, p. 261. (The Loeb Classical Library.)

[41] Liddell and Scott's Lexicon, s.v.

[42] *Op. cit.*, Vol. III, p. 25.

That βάπτεσθαι means that the bronze received a color in dipping, as Frazer interprets it, seems at first glance very likely from the ways in which the verb βάπτω is often used by ancient authors other than Pausanias. This verb is often used to signify the dyeing of cloth and other materials, and is even used in early Greek alchemical works with reference to the gilding or silvering of metals. Moreover, the closely related verb καταβαπτω is used in the Leyden Papyrus X,[43] probably composed at a time not far distant from that of Pausanias, to signify the dipping of base silver alloys in corrosive liquids for the purpose of changing or improving the surface color of such alloys. However, in spite of numerous extant examples of these closely related uses, there appears no actual example in the works of other authors of the use of the verb βάπτω itself to denote the coloring of bronze or copper. From the ways in which βάπτω is used by these other ancient authors no certain indication therefore exists that the word βάπτεσθαι in the statement of Pausanias really means that the bronze received its color by dipping.

In translating βάπτεσθαι in the way he did, Frazer appears to have been guided to a considerable extent by Göttling,[44] who not only interpreted the statement of Pausanias as meaning that the bronze received a color by being quenched in the water of Peirene, but suggested also that this coloration was the result of the deposition of ocherous material from this water on the metal by such treatment, a suggestion based upon his observation while visiting Corinth of deposits of ocherous material apparently formed by this water. That the bronze could have been permanently or satisfactorily colored in such a way is highly improbable. At the time the present writer visited Corinth no ocherous deposits formed by the water of Peirene were to be seen. Moreover, in experiments in which samples of this water were evaporated to dryness no colored residues were obtained, only residues of white salts. That the water could have changed radically in composition between the times of these two visits is also highly improbable. The suggestion of Göttling appears, therefore, to be unsound.

Though the coloration of the bronze by the mere evaporation of the water on the hot metal with the consequent deposition of a colored residue seems to be a possibility that must be excluded, the coloration of the bronze by material formed by some chemical interaction between the hot metal and the salts dissolved in the water is a possibility still to be considered. In connection with some remarks on the fame of Corinth as a center of bronze manufacture, Davies [45] cites an opinion of Desch that the composition of water in which bronze is quenched is without effec on the color of the metal, and this is certainly what is to be expected from general chemical considerations. Nevertheless, i order to settle the question decisively in respect to the water o the Fountain Peirene, a considerable number of systematic experiments were performed by the writer. In these experiments a piece of pure copper, a piece of ordinary tin bronze of modern manufacture, and a piece of clean bronze taken from the core of the Corinthian strigil handle previously described were heated to various elevated temperatures, either in a bare flame or without contact with a flame, and were then plunged, either immediately or after various definite times, into water from the Fountain Peirene, into distilled water, or into dilute sodium chloride solutions of various concentrations. The detailed results of the many experiments need not be related here since, in general, it was found that any coloration imparted to the metals was solely the result of oxide films formed during the heating of these metals and not to any chemical interaction between the hot metals and the liquids into which they were plunged. If these liquids had any effect at all it was that of suddenly stopping the oxidation of the surface of the metal at a definite stage by cooling it so that a particular coloration obtained on heating did not afterwards change, as might happen on allowing the metal to cool slowly in air. However, the water of Peirene was of no more value for this purpose than any other of the liquids that were tried, nor did the piece of Corinthian bronze behave very differently from the piece of copper or the piece of modern bronze. Moreover, the colors imparted to the Corinthian bronze or to the other metals were not especially pleasing or striking, and were decidedly not attractive when these metals were heated to a high temperature, as the account of Pausanias clearly indicates was the practice in the treatment of bronze with the water of Peirene. When the metals were heated to a high temperature for some time so as to form a black coating or scale of oxidation products on their surfaces and were then plunged into this water much of this coating or scale cracked off so as to expose a slightly oxidized surface of nearly the same color as the bare metal. This may have been the purpose of the process mentioned by Pausanias, but here again the water of Peirene was found to be of no more value for the removal of scale than any other of the liquids that were tried, nor did the specimen of Corinthian bronze behave in any distinctive or peculiar way. Of course it may be supposed that a method for imparting a coloration to bronze by roasting the metal in air and cooling in water was first developed at Corinth, or that the technique of restoring something like the original color of the metal by the removal of scale by sudden cooling in water was first practiced at Corinth, and that as a consequence it came to be falsely and traditionally believed that the water at Corinth was especially suitable for use in the one or the other of these technical procedures. However this does not seem likely since the superficial coloration of bronze on heating in air, and the dropping off of scale on sudden cooling with water are such simple and obvious effects that they must have been observed by experienced ancient bronze workers at many localities long before the manufacture of bronze began at Corinth. Hence it does not seem possible that any special quality could have been anciently ascribed to the water of Peirene by reason of any changes of color that resulted from merely heating bronze to elevated temperatures and then plunging it into this water.

---

[43] Recipes 19 and 20 of the text of Berthelot published in his *Archéologie et Histoire des Sciences*, Paris, 1906, pp. 268–306.

[44] *Archäologische Zeitung*, 2, 328 (1844). *Cf.* Frazer, *op. cit.*, Vol. III, p. 24.

---

[45] *Roman Mines in Europe*, Oxford, 1935, p. 252.

Apparently the only possible way by which the water of the Fountain Peirene could have been used to impart a distinctive coloration to bronze was by the accidental or intentional concentration of the water before the hot metal was dipped into it so that an appreciable residue of salts was left on the surface of the metal through evaporation of the film of water that remained after dipping, these salts then causing the formation, after some time, of a green coating of copper salts from the corrosive action of the chlorides contained in the residue of saline material. If water from the fountain had been kept in containers in the shops of metal workers, and these containers had been kept filled by the addition of water and not by being emptied out and refilled each time more water was needed, the necessary concentration of the water could have occurred in such containers both by spontaneous evaporation and by the action of the hot metal dipped into the water. In the experiments it was found that solutions of sodium chloride of about the same chloride concentration as concentrated water from Peirene had less effect than the concentrated water, probably because the saline deposit left on the surface of the metal by the water was much more hygroscopic by reason of the presence in it of calcium or magnesium chloride. The piece of Corinthian bronze was found to acquire a superficial greenish coating much more readily than the piece of pure copper or the piece of modern bronze, probably because the ancient metal was much more heterogeneous. However, the development of a greenish coating or rough patination by the chemical action of salts left on the surface of the bronze was found to be a slow and uncertain process at the best, one that often took several weeks. Though the water of the Fountain Peirene apparently could have been used to impart a distinctive coloration to bronze in a way impossible by a similar use of water from many other ancient sources, it seems very unlikely that such a slow and troublesome process was actually used in the treatment of bronze at Corinth, and still more unlikely from the wording of his statement that Pausanias is alluding to any such process.

A further serious objection to the translation of *βάπτεσθαι* in the way suggested by Göttling and followed by Frazer is that there was neither need nor desire to impart any coloration whatsoever to many kinds of bronze objects, especially those objects intended for useful purposes. Though the same objection cannot be said to hold generally for ornamental objects in bronze, it may, nevertheless, be valid for the ornamental objects made at Corinth since the statement of certain ancient authors [46] respecting the appearance of objects of Corinthian bronze certainly seem to indicate, aside from any question as to their appearance being the result of the presence of gold and silver in the alloy, that such objects were admired by reason of the beauty of the bare metal not by reason of any artificial coloration or patination imparted to the surface of the bronze.

From all the evidence taken together the translation of *βάπτεσθαι* as meaning that the bronze received a color on being dipped appears, therefore, to be incorrect, and to lead to wrong interpretation of the statement which Pausanias makes about the treatment of bronze with the water of the Fountain Peirene.

[46] *E.g.*, Pliny, *H.N.*, XXXIV, 6–8; Plutarch, *De Pythiae Oraculis*, 395 B–D.

That βάπτεσθαι means, in accordance with the translation of ones, that the bronze was tempered by being dipped in the ater may be said to receive considerable support from the way n which βάπτω and related words are often used by ancient riters other than Pausanias. βάπτω is clearly used to signify he dipping of red-hot steel in water for the purpose of hardenng this metal, and βαφή is used to denote this dipping or hardenng process. By extension of meaning βαφή even came to mean he hardened edge imparted to iron or steel tools by such a process. Though there can be no doubt that these words were used with reference to the hardening or tempering of steel by a dipping process, the evidence that βάπτω and related words were used with reference to the hardening or tempering of bronze by some similar process is by no means so clear or abundant from the works of these writers. Aeschylus,[47] indeed, uses the expression χαλκοῦ βαφάς, which may be interpreted as meaning the hardening or tempering of bronze or copper, but this expression is considered by some editors or commentators as being merely poetical for σιδήρου βαφάς, the hardening or tempering of iron or steel, though their reasons for supposing that the expression which appears in the established text is merely poetical cannot be said to be substantial. However, Pollux[48] in noting the use of βάψις in place of βαφή quotes Antipho Sophisto as using the expression βάψιν χαλκοῦ καὶ σιδήρου, which may be interpreted as meaning the hardening or tempering of bronze and of steel, or the dipping of the metals for such a purpose. The fact that the names of the two metals are thus linked together seems to indicate, at least, that some similar dipping process was used for both.

Late ancient and medieval grammarians or commentators on the works of Classical Greek authors speak explicitly of the former existence of a process for hardening bronze by dipping. For example, in commenting on some verses of Hesiod[49] in which the poet alludes to the use of bronze for various purposes before iron came into use, Joannes Tzetzes remarks,[50]

---

[47] *Agamemnon*, v. 612 (or 617).

[48] *Onomasticon*, VII, 169.

[49] *Opera et Dies*, v. 150 ff.

[50] Text as quoted by Rossignol, *Les Metaux dans L'Antiquité*, Paris, 1863, p. 239. Rossignol also quotes two other passages of very similar import.

Χαλκοῖς τὸ παλαιὸν καὶ ὅπλοις, καὶ ξίφεσι, καὶ γεωργικοῖς ἐργαλείοις ἐχρῶν βαφῇ τινι ταῦτα στομοῦντες. Ἀπολλυμένης δὲ τῆς στομούσης βαφῆς τὸν χαλκ χρώμεθα τῷ σιδήρῳ. (Anciently they used bronze for armor, swords and agricultural implements, hardening this by a certain dipping process, but the dipping process having been lost, we use iron.

Such statements indicate that there existed at a late period a general belief that the ancients once had a process for hardening bronze or copper similar to the one used for hardening steel, and that this supposedly valuable process had become a lost art. Even at the present day the same notion seems to be widely current. Apparently the earliest clear statement of this belief appears in the *De Pythiae Oraculis* of Plutarch written about 100 A.D. In this work Plutarch has Philinus quote a foreign visitor at Delphi as asking,[51] Ἀρ' οὖν κρᾶσίς τις ἦν καὶ φάρμαξις τῶν πάλαι τεχνιτῶν περὶ τὸν χαλκόν, ὥσπερ ἡ λεγομένη τῶν ξιφῶν στόμωσις ἧς ἐκλειπούσης ἐκεχειρίαν ἔσχεν ἔργων πολεμικῶν ὁ χαλκός; (Was there, then, some process of alloying and treating used by the artisans of early times for bronze, something like what is called the tempering of swords, on the disappearance of which bronze came to have a respite from employment in war?). In spite of the fact that Plutarch writes as though the process of hardening bronze was in his day considered to be an art which had fallen into disuse or had been forgotten, and in spite of the fact that Pausanias lived a little later than Plutarch, it is nevertheless not impossible that Pausanias believed from his own opinions, or from what he was told by others at Corinth, that the hot bronze was dipped into the water of Peirene for the purpose of hardening the metal. Since his phraseology indicates that he had no first-hand knowledge of the process, it seems rather more likely that he may have been led to such a conclusion merely by what the bronze workers or other people told him. Moreover, because he was a more or less unknown visitor from distant parts, they may even have deliberately deceived him as to the real purpose of the process, especially since the important details of technical processes appear to have been guarded as trade secrets by the artisans of ancient times. Thus in one way or the other Pausanias may have come to believe that the purpose of the process was to harden the bronze. Consequently in his statement about it the word ττεσθαι may be intended to convey just this meaning, and it is erefore by no means unlikely that the word should now be so terpreted in translation.

---

[51] Both text and translation essentially as given by Babbitt, *Plutarch's Moralia*. Cambridge, 1936, Vol. V, pp. 262--263 (The Loeb Classical Library).

However, if Pausanias believed that the bronze was hardened this process he was clearly wrong since the truth is that onze, unlike steel, cannot be hardened by plunging the red-hot etal into cold water. The effect of such treatment on the physal properties of ordinary bronze may be very slight, or it may nd to soften the metal somewhat, though if the bronze contains high enough proportion of tin this treatment may have one ry important effect on the working qualities of the metal. hen bronze containing over ten and up to about fifteen per cent tin is heated to above 650° C. for a sufficient time and is then ddenly quenched in cold water, this bronze then becomes ctile enough to be worked by hammering while cold, whereas st bronze of this composition, and in general any bronze of this mposition not so treated, cannot be shaped by hammering hile cold because it is too brittle. That the temperature atained in the process mentioned by Pausanias was high enough nd that it was continued long enough seems to be indicated by is words διάπυρον καὶ θερμὸν. Moreover, as was shown by the anatical results given at the very beginning of this paper, high-in bronze was certainly produced at Corinth at an early period, nd by reason of the persistence of local technical tradition in ncient times, it may have been produced there at the time of ausanias, or not very long before his time. In order to work uch bronze by hammering it while cold the very treatment that ausanias mentions would have been almost necessary. This reatment would have been virtually a necessity in the manufacure of very thin objects of high-tin bronze since these could not ave been satisfactorily forged while hot because of their tendncy to cool down too rapidly and because of their tendency to ecome oxidized too much by repeated heating to a high temerature. Some definite indication that the same treatment may ave been used in the manufacture of even moderately thin objects was found, it may be recalled, in the course of the microscopic examination of the metallic core of a rather typical object of this sort, a strigil handle found at Corinth and presumably made there. Even if high-tin bronze was not being produced or worked at Corinth at the time of the visit of Pausanias and had not been produced or worked there for a long period before his visit, which is by no means unlikely in view of the probable scarcity of tin in the Mediterranean world in Roman Imperial times, this quenching process, because it was a local practice of long standing, may, nevertheless, have still been used at Corinth for the working of ordinary bronze even when there was no longer much real advantage in using it. On the whole, therefore, the process which Pausanias describes probably had as its real purpose, originally or possibly even at the time of his visit to Corinth, the softening of high-tin bronze so that this metal could be worked while cold. It may even be that Pausanias was aware of the real purpose of the process, and that he meant to convey this meaning in the word βάπτεσθαι, though this seems very improbable since no such meaning was ever attached to the verb βάπτω or to related words by any other ancient author.

Regardless of whether Pausanias wrongly believed that the bronze was hardened by the dipping process he mentions or whether he rightly believed that the bronze was softened by this process, the way in which Jones has translated βάπτεσθαι is satisfactory because the English verb "to temper" as applied to metals may be taken to signify that either a hardening or a reduction of hardness takes place. Hence the translation of Jones provides a happily ambiguous English equivalent of a still more ambiguous Greek word.

An apparent difficulty in the way of accepting as wholly satisfactory the general interpretation of the statement of Pausanias which results from this translation of βάπτεσθαι is that this interpretation does not explain why the water of the Fountain Peirene was considered especially suitable for the treatment of the bronze, as the wording of the second phrase of his statement seems to imply. The fact of the matter is, of course, that the composition of a natural water can have no effect on the physical properties of the bronze quenched in it. In other words from the standpoint of the actual results of this operation there was no reason why the water of Peirene was to be preferred to that from any other spring, or to that from any river or well. Nevertheless, ample evidence exists to show that the ancients really did believe that the character of the water had an important influence on the quality of the metal quenched in it. This erroneous belief appears clearly in some remarks which Pliny makes about the manufacture of steel. After discussing the different kinds of iron or steel and the reasons for the differences Pliny goes on to say,[52] *Summa autem differentia in aqua cui subinde candens immergitur. Haec alibi atque alibi utilior nobilitavit loca gloria ferri, sicuti Bilbilim in Hispania et Turiassonem, Comum in Italia, cum ferraria metalla in iis locis non sint.* (The chief difference, however, arises from the nature of the water in which the glowing metal is plunged from time to time. This, in some places better for the purpose than in others, has quite ennobled some localities famous for their iron such as Bilbilis and Turiasso in Spain, and Comum in Italy, even though there are no iron mines at these localities.) Since a special value for the quenching of steel was thus erroneously attributed to the water of some places, it seems very likely by analogy that a special value for the quenching of bronze was likewise erroneously attributed to the water of certain places, of which Corinth appears to be the only one now known. Moreover, it will be noted as an interesting and perhaps significant parallel that just as there were no iron mines at the places where the water was reputed to be superior for the quenching of steel so also were there no copper mines at Corinth where the water was apparently considered superior for the quenching of bronze.[53]

A clear indication that special properties were anciently ascribed to certain materials for the quenching of bronze in order to render the metal soft enough for working is given by one of the recipes of the Leyden Papyrus X. According to the text published by Berthelot[54] this recipe, which is No. 67, reads as follows:

Χαλκοῦ μάλαξις

Πυρώσας ἐντίθει εἰς κόπρον ὀρνιθείων, καὶ ὅταν ψυγῇ, ἔλαυναι.

---

[52] *H.N.*, XXXIV, 144. Text of Bailey, The Elder Pliny's Chapters on Chemical Subjects, Part II, London, 1932, p. 58.

[53] The absence of copper deposits or any trace of ancient copper mining at Corinth or in its immediate vicinity has been established by careful search. According to Davies (*Roman Mines in Europe*, Oxford, 1935, p. 253) the nearest locality for copper minerals is at the top of the pass between Mycene and Phlius, and even here the amounts to be seen are very small. No ancient workings have been found at this locality. That copper ore was actually smelted at Corinth appears, however, to be demonstrated by the finding by Davies of a piece of copper slag at the site. This was identified as such by chemical analysis.

[54] *Archéologie et Histoire des Sciences*, Paris 1906, p. 292.

This may be translated in the following way:

Softening of Bronze

Fire it, place it in bird dung, and when it cools, forge it.

This recipe clearly indicates a belief in the efficacy of a particular material for cooling bronze quickly after heating it to a high temperature in a process for softening the metal so as to be able to work it easily while cold. No real basis exists for preferring this material to water. Whether bird dung was actually the material employed is perhaps somewhat uncertain since the Greek term, which has here been taken literally, may in fact be a technical term for a very different material, the use of cryptic technical terms in early technical recipes being by no means uncommon.

From all this it seems reasonably certain, then, that the ancients could have believed that the water of Peirene had a special value for the treatment of bronze when as a matter of fact it had no such value. The question remains as to how such a false belief could have arisen.

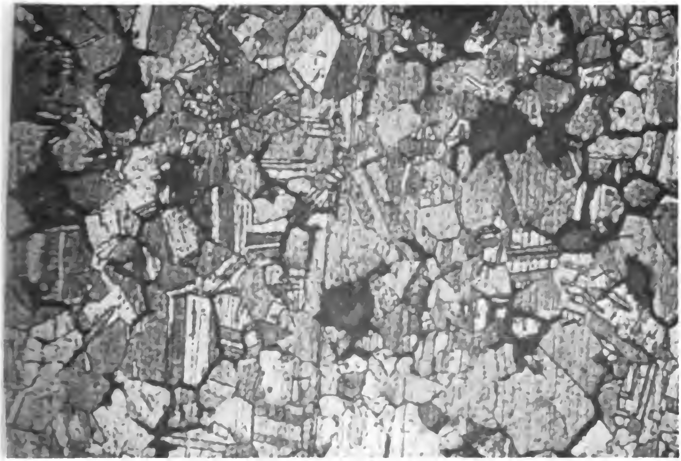
It may be of course that the reputation of the water of Peirene for the treatment of bronze arose merely because this fountain was a never failing and abundant source of water available at all times for metallurgical work, which is a good deal more than can be said for many other water supplies in Greece. But this does not seem sufficient to account for the high reputation which the water apparently had for this particular purpose. Perhaps the general reputation of the water and the fame of Corinth for fine work in bronze were sufficient to lend credence to the belief that the water of this noted fountain used in the treatment of the bronze had something to do with the quality of the metal. What seems somewhat more likely, however, is that some bronze worker at Corinth discovered accidentally that high-tin bronze could be forged in the cold state after a certain empirical treatment which included the quenching of such bronze in the water of Peirene after the metal had been heated to a given high temperature for a given time, the success of the process as a whole being wrongly ascribed to the nature of the water instead of to the method of heat treatment. Since the success of the quenching process in the hardening of steel in ancient times appears to have been generally ascribed to the nature of the vater in which the metal was quenched, it is easy to see how, by he same kind of false reasoning, the success of the process for treating bronze could also have been ascribed to the nature of the water. Even if it had become recognized sooner or later that the nature of the water had nothing to do with the success of the annealing treatment for bronze, it is unlikely for obvious reasons that the bronze workers who used the method would have allowed the truth to become generally known. Possibly the discovery and exclusive use of a method for working high-tin bronze in the cold state may have been one of the reasons why Corinth became a famous center of bronze manufacture in ancient times.

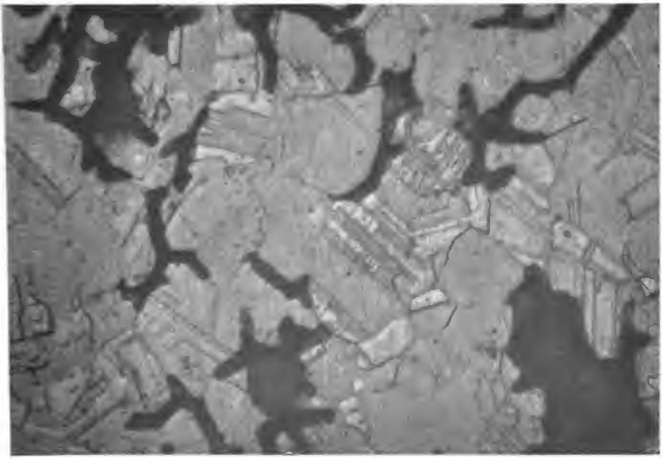
This long discussion of the brief remark which Pausanias makes in regard to the treatment of bronze at Corinth with the water of the Fountain Peirene serves to indicate the value of chemical and technological facts and experiments for the interpretation of certain passages in the writings of ancient authors which are not capable of being adequately explained from historical and philological evidence alone.

ACKNOWLEDGMENT

The author again expresses his thanks to the American Philosophical Society for the grant made to him from the Penrose Fund in 1936 since important parts of the work here described were financed from this grant.

PLATE I

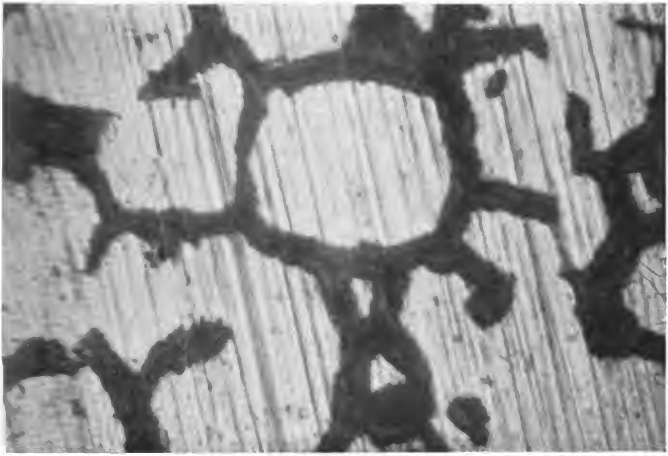


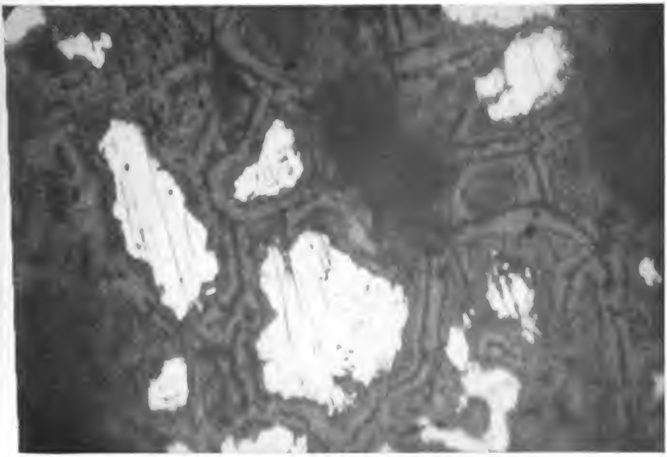


Top Figure. Photomicrograph of best preserved part of metallic core of a corroded strigil handle found at Corinth showing homogenized and twinned grains of variable but generally small size, a few corroded grains, and some intergranular corrosion. Magnification 500 ×.

Bottom Figure. A view at high magnification in the same region as in the upper figure showing in particular the presence of numerous slip lines on some of the grains and the existence of highly corroded metal in the presence of uncorroded metal. Magnification 2150 ×.

PLATE II

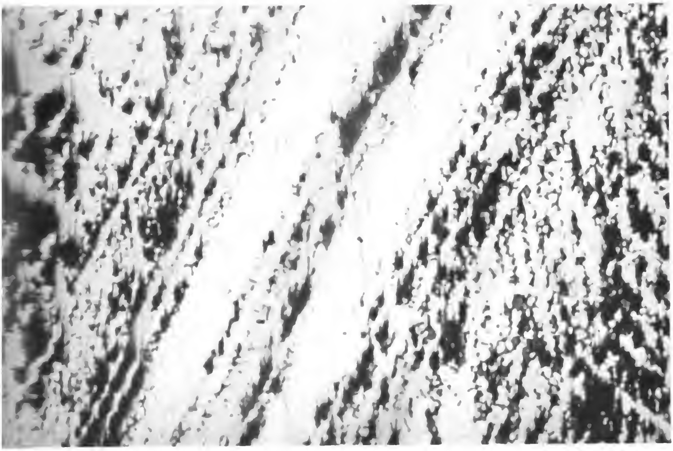


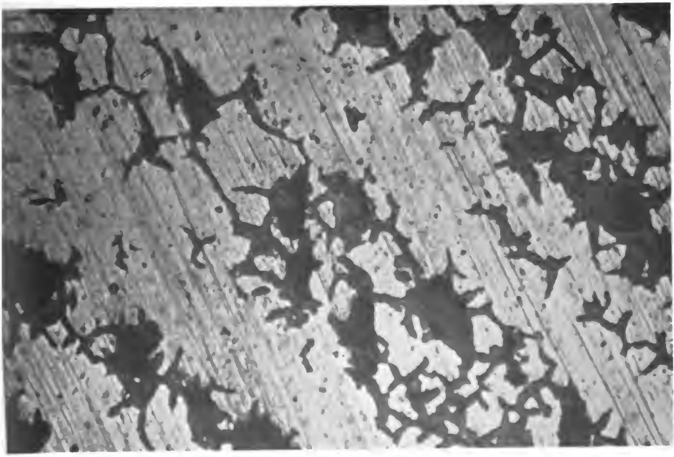


Top Figure. Photomicrograph of selected region in metallic core of strigil handle showing grains of metal completely surrounded by intergranular corrosion products. Magnification 2150 ×.

Bottom Figure. Photomicrograph taken near edge of metallic core of strigil handle showing progressive corrosion of grains of metal inwards from the grain boundaries. Magnification 2150 ×.

PLATE III



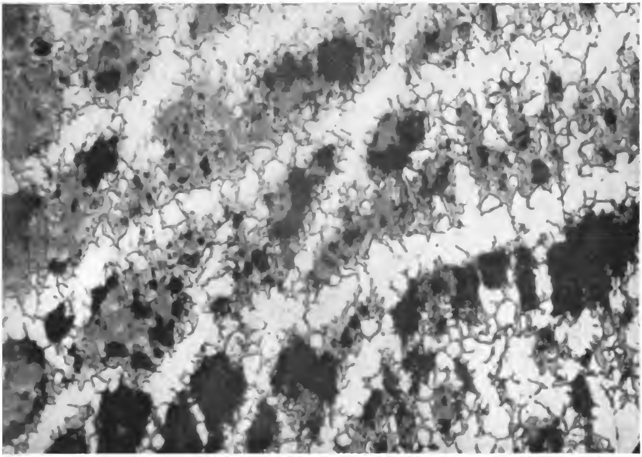


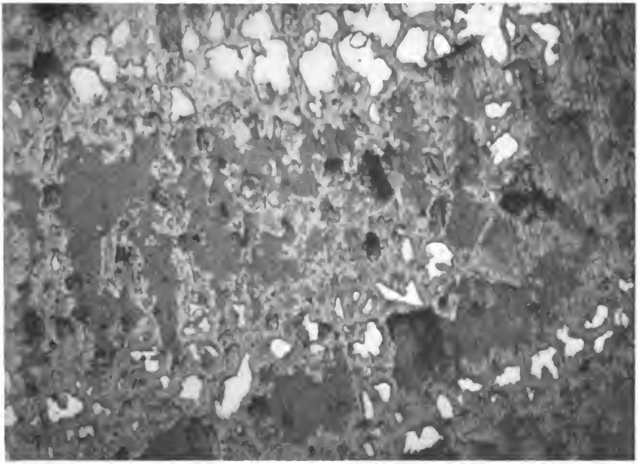
Top Figure. View at low magnification of the general pattern of the intergranular and granular corrosion within the metallic core of the strigil handle. Magnification 120 ×.

Bottom Figure. View at higher magnification of the alternate layers of corroded and uncorroded metal in the metallic core. Magnification 500 ×.

Copyrighted material

PLATE IV

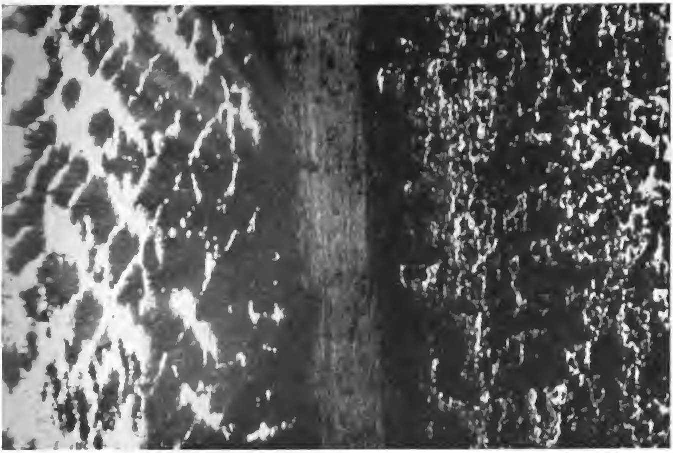


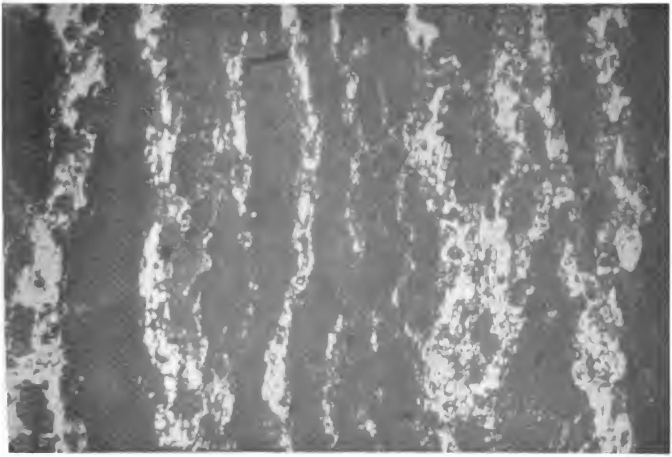


Top Figure. Photomicrograph of extensively corroded metal at edge of metallic re showing survival of cross layers of uncorroded metal. Magnification 250 ×.

Bottom Figure. A view of a still more extensively corroded region at edge of etallic core showing the confused and heterogeneous structure. Magnification 500 ×.

PLATE V





Top Figure. Photomicrograph showing a cross section at low magnification of one side of the metallic core of the strigil handle, most of the patina on this side, and a transition layer between the two. The patina is at the right. Magnification 120 ×.

Bottom Figure. A view in cross section of the patina on the strigil handle showing the pronounced banded structure. Magnification 250 ×.

ABORIGINAL AUSTRALIAN STRING FIGURES

DANIEL SUTHERLAND DAVIDSON

Assistant-Professor of Anthropology, University of Pennsylvania

Preface

The material presented in this report was collected in North Australia in 1930 on behalf of the Museum of the University of Pennsylvania, in Western Australia, South Australia and Victoria in 1930 and 1931 under a Fellowship grant from the Social Science Research Council, and in Western Australia during 1938 and 1939 under the auspices of the University of Pennsylvania, the American Philosophical Society (Penrose Fund) and the Carnegie Corporation of New York. To all these institutions the author wishes to express his very sincere gratitude for their interest and support.

TABLE OF CONTENTS

PAGE
Preface 763
Introduction 765
Australian String Figures 771
Antiquity of String Figures in Australia 777
Melanesian Derivation of Australian String Figures 783
Basis for Comparative Studies of String Figures 786
Nomenclature 793

String Figures Formed from First Position

1. Two Dogs (Fig. 9) 799
2. A Water Hole (Fig. 10) 800
3. Two Water Holes (Fig. 11) 801
4. A Dead Man in a Tree (Figs. 12–13) 802
5. Phallus (Fig. 14) 803
6. Dilly Bag (Figs. 15–17) 804
7. Baby Being Born (Figs. 18–19) 806
8. Big River Bat (Figs. 20–21) 807
9. Blackfellow (Fig. 22) 809
10. Lily (Figs. 23–26) 811
11. Headdress (Fig. 27) 813
12. Hut (Fig. 28) 814
13. Emu Foot (Fig. 29) 815
14. Emu (Fig. 30) 816

PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY,
VOL. 84, NO. 6, AUGUST, 1941

[illegible] OSOPHICAL SOC
AUG 28 1941

STRING FIGURES FORMED FROM OPENING A

15. Two Kangaroos (Fig. 31) .... [illegible]
16. Two Turtles on a Log (Fig. 32) .... [illegible]
17. Kangaroo Trail (Fig. 33) .... [illegible]
18. Dead Kangaroos (Figs. 34–37) .... [illegible]
19. Big Mob of Blackfellows (Fig. 38) 822
20. Armband (Fig. 39) 824
21. Fish Spear (Fig. 40) 826
22. Vulva (Fig. 41) 828
23. Spear (Fig. 42) 828
24. Bed (Fig. 43) 829
25. Bunk (Fig. 44) 831
26. Eaglehawk (Fig. 45) 831
27. Big Pig (Figs. 46–49) 833
28. Dilly Bag (Fig. 50) 835
29. Turtle (Fig. 51) 836
30. Turtle (Fig. 52) 838

STRING FIGURES FORMED FROM OPENING A, MOVEMENT Z

Movement Z (Figs. 53–54) 839
31. Two Blackfellows (Fig. 55) 841
32. Bamboo (Fig. 56) 842
33. Rain (Fig. 57) 842
34. Devil-devil's Anus (Fig. 58) 844
35. Crocodile (Fig. 59) 844
36. Crocodile's Eye (Fig. 60) 845
37. Gate (Fig. 60) 847

STRING FIGURES FORMED FROM OPENING B

38. Men Coming to a Fight (Figs. 61–63) 847
39. Goanna (Figs. 64–66) 849
40. Bunk (Fig. 64) 851
41. Python (Figs. 67–69) 852
42. Boomerang (Fig. 70) 853
43. Canoe (Figs. 71–72) 854
44. Water (Fig. 73) 855
45. Blackfellow Steals a Lubra (Figs. 74–77) 855
46. Boy After Lubra Runs Away (Figs. 78–79) 858

FLYING FOX SERIES

47. Canoe (Flying Fox Series) (Fig. 80) 860
48. First Flying Fox (Flying Fox Series) (Fig. 81) 860
49. Second Flying Fox (Flying Fox Series) (Fig. 82) 861
50. Third Flying Fox (Flying Fox Series) (Fig. 83) 861
51. Fourth Flying Fox (Flying Fox Series) (Fig. 84) 862

SERIES (NINE FIGURES)

52. (a) Boomerang (Fig. 85) 863
53. (b) Python (Fig. 86) 863
54. (c) Inverted Boomerang (Fig. 87) 864
55. (d) Unnamed Figure (Fig. 88) 865
56. (e) Bandicoot (Fig. 89) 865
57. (f) Unnamed Figure (Fig. 90) 866
58. (g) Unnamed Figure (Fig. 91) 867
59. (h) Unnamed Figure (Fig. 92) 867
60. (i) Water (Fig. 93) 867

String Figures Formed from Opening C

PAGE
61. Four Blackfellows (Figs. 94–95) 868
62. Canoe (Fig. 96) 870
63. Net (Fig. 97) 871
64. Turtle (Fig. 98) 872

String Figures Formed from Elbow Opening

65. Tomahawk (Fig. 99) 873
66. Lightning (Fig. 100) 874
67. Axe (Fig. 101) 876

String Figures Formed from Odd Openings

68. House (Figs. 102–103) 877

String Tricks

69. "Cutting Off the Head" (Fig. 104) 879
70. "Hanging" (Fig. 105) 879
71. Bullocky (Fig. 106) 880
72. "Cutting Off the Fingers" (Figs. 107–109) 881
73. Cutting Off the Nose (Figs. 110–111) 882
74. "Threading the Needle" (Fig. 112) 884

Appendices

A. String Figures from Arnhem Land, North Australia 885
B. String Figures from Hermannsburg, Central Australia 886
C. String Figures from the Coorong, South Australia 888
D. String Figures from the Diamantina District, South Australia 889
E. Comparative Chart of Australian String Figure Distributions 890
F. Cup and Saucer (Fig. 113) 900
Bibliography 900

INTRODUCTION

Of common appearance throughout the world are games or pastimes which consist of intertwining on the figures a closed loop of string for the purpose of producing some pleasing pattern or orderly arrangement. If for want of a better name we may speak of them collectively as string games, three main classes can be recognized: (1) String tricks and catches, (2) cat's cradle, and (3) string figures. All are similar in general principle and presumably are related historically, but each is characterized by such definite peculiarities that a ready distinction can be made.

String Tricks and Catches

String tricks and catches, as the terms imply, find their effectiveness in the element of surprize; for the tricks, the strings from some apparently complex arrangement are suddenly released to restore the original loop; for the catches, the strings when released from a more or less simple pattern unexpectedly lock a finger or hand. Of not uncommon appearance throughout the world both tricks and catches are extremely simple and readily learned. They seem to be fairly popular as a class among seafarers, traders and other itinerant whites and it is possible that certain examples have been learned in the past by such groups at various places and ultimately brought to many localities of their present worldwide distribution. At the same time it is not unlikely that their appearance in other areas is the result of diffusion under strictly aboriginal conditions. Whatever may be the truth of the matter the question of their individual origins is most obscure and cannot be discussed profitably in the light of present knowledge.

Cat's Cradle

Cat's cradle, well known to western Europeans, consists of a series of simple figures executed by two players who alternate in taking a design off the hands of the other and, by a simple manipulation, remaking it into the succeeding pattern of the series. It is distinguishable from string figures in a number of ways. The opening, or initial arrangement of the strings on the hands, is peculiar to itself and not associated with other types of figures. The procedure also is distinctive both in the manner whereby the strings are lifted off the hands of the other player and in the simple manipulation to form the next extension or resulting design. Finally, the patterns, themselves, are dissimilar in principle in that the original loop generally can be restored without difficulty simply by dropping the strings at any step in construction.

In addition to these most elementary technical features cat's cradle is characterized by a most remarkable uniformity of procedure wherever it is played, although all the patterns in the series are not reported everywhere.[1] In view of this general similarity it would seem that the number of possible designs obtainable from the cat's cradle opening is strictly limited to those known, or that the peoples who have made these patterns for generations either have not been interested in experimenting for additional patterns or have been unsuccessful in their quest. Be this as it may, we find cat's cradle today institutionalized everywhere to the point of monotony and unaccompanied by any imaginative interest.

---

[1] For a discussion of these details see Jayne, pp. xii, 324.

Before the period of European expansion cat's cradle seems to have been confined to eastern Asia, its apparent place or origin, the adjacent islands of the East Indies, Korea and Japan, and western Europe, where its introduction is believed to have accompanied the early tea trade. It seems to have been unknown throughout the remainder of the world until European colonists penetrated the other continents and brought this pastime with them. Today it is apt to appear almost anywhere among native peoples who often are unaware of its European derivation.

STRING FIGURES

String figures, in the widest sense of the term, properly include both string tricks, catches and cat's cradle. However, for purposes of classification and discussion it is advantageous to recognize a basis of differentiation as determined by both technical and distributional considerations. As will be brought out more fully below, string figures are characterized by a variety of openings, movements and methods of extension exclusively their own. In direct contrast to cat's cradle the final patterns, with a few exceptions, are definitely locked in the sense that the strings are interwined in such a manner that the return to the original loop cannot be accomplished just by dropping the strings, but usually must be attained by some specific release.

The manipulation of most string figures is performed entirely by the fingers and hands of one person, but not infrequently the teeth, neck, elbows, knees or toes and the services of two or even three individuals may be required to produce the finished pattern. Although in most figures the objective is a single design, there are examples in which the finished pattern is the optional starting point for one or more different figures. In some cases series are performed by a single player, in others the procedure may require two manipulators who either alternate in making the constituent progressive figures or produce them by joint effort.

String figures also are characterized by a relative complexity of procedure and intricacy of pattern. In remarkable contrast to the monotonous and strictly limited cat's cradle, which everywhere is performed perfunctorily and equally well by the stupid and clever, string figures are featured by such a variety of opening arrangements, types of manipulation and methods of extension that strict attention to detail is necessary throughout the procedure and careful memorization of the proper continuity required. Indeed the steps in construction in some figures are so complex that a correctly attained final result can be considered an accomplishment of no mean order. Hence, in those areas where string figures are well established there usually is a sufficient variety to attract some interest from all the intelligence levels of the population. For the dolts there is the simple type of figure; for the brilliant, a potentially unlimited source for mental exercise and amusement; for the experimentalist, a fruitful field for development with the discovery of new and intricate patterns as a reward; for the artist, an opportunity for keen personal satisfaction in the meticulous execution of difficult figures; for the skilled entertainer, an extensive repertoire with which to delight an audience.

The distribution of string figures is curious and difficult to explain satisfactorily at the present time. Of widespread appearance in North America, South America, Africa, Oceania and Australia, they are lacking strangely enough in Europe and Asia except for four peripheral areas of relatively minor extent, Western Europe where a very few figures of doubtful origin and antiquity are found; India for which several figures and tricks are reported as well as some peculiar Brahminic figures apparently of recent and possibly independent origin: the Philippines, where some figures are known to mountain tribes, whereas cat's cradle is found in the lowlands; and northeastern Siberia, where a great many examples have been collected. Thus, in recent centuries string figures seem to have been found in Asia only on or beyond the northern, southeastern and southwestern boundaries of the distribution of cat's cradle.

The virtually exclusive distributions of cat's cradle and string figures are puzzling, to say the least, for the peculiar situation in Asia suggests that simple string figures formerly were more widespread and have been replaced in part by equally simple cat's cradle. Why the latter with its strictly limited possibilities and relatively uninteresting manipulation should ave had more appeal than even the simplest string figures, hich in all other parts of the world have provided a basis for ich and complex developments, is difficult to understand, for ost Asiatics for many generations have evinced keen interest nd delight in games and puzzles which require mental exercise.

The situation, however, is not easy to assess, for a former idespread appearance, of which the modern Asiatic occurrences may be vestiges, possibly antedated the beginnings of the great civilizations of the continent and the intellectual interests accompanying them. More detailed comparative studies must be made in many areas before we can hope to secure definite clues to such a question. However, it is interesting to note that a general apathy toward string figures is found even today not only throughout Asia but in Europe as well, for in spite of the fact that several hundreds of examples from various parts of the world are now readily available in the literature, not even the mentally curious have shown any extensive interest in them or in the theoretical technical aspects of their construction.

One reason why string figures have attracted so little attention in modern civilization undoubtedly is the recency of our knowledge of them. Practically all of our detailed information has been collected since the beginning of the present century, a large proportion of it within the past twenty years. Previously only a few Europeans had ever heard of string figures and their very existence among native peoples apparently was unknown in Europe until about a century ago.

Although more extensive research may bring to light earlier references, the first mention of string figures known to the author is that of Dieffenbach who noted their presence in New Zealand in 1843.[2] His observations actually may have been antedated by those of Thomas Petrie who some time between 1840 and 1850 noticed the Australian natives in the Brisbane area playing *warrowarro,* a term they subsequently applied to the whiteman's fences because of their resemblance to string figure patterns.[3] In 1844 Cawthorne called attention to the custom for the Adelaide tribe[4] and in 1845 Eyre added his observation for the Southeast.[5]

---

[2] Dieffenbach, Vol. II, p. 32.

[3] Roth, 1903, p. 583.

[4] Cawthorne, p. 14.

[5] Eyre, Vol. II, p. 227. Another early reference (1857) is Bunce, p. 75 for Victoria.

In all other parts of the world the discovery of string figures by Europeans seems to have been reported only at a later date. They were noted among the Eskimo in 1864 and in Polynesia in 1876.[6] Elsewhere their presence was recorded in still more recent times and even today there are many tribes for which no information, either positive or negative, is available. Although it seems likely that they are not known to all native peoples there can be little doubt that investigation would reveal their presence among many of them.

Of the early writers none attempted to describe the method of making the string figures they saw, nor did they illustrate the patterns they witnessed. Their information, therefore, is of little actual value except as historical records, for it throws no light on the many technical problems of string figure development and diffusion. Even at the present time few field workers take the trouble to record the methods of manipulation in the areas they investigate.

The first scientific interest in string figures dates from 1879 when Tylor, in a general discussion of games, drew attention to their ethnological significance.[7] In 1881 Klutschat illustrated the finished designs of three Eskimo figures,[8] but the first serious interest in the field was shown by Boas who in 1888 made the initial effort to describe the intricate manipulation employed by the Eskimo.[9] There was no immediate response from others to record procedure, but Tenicheff (1898) and Kroeber (1899) did illustrate the patterns of other Eskimo groups.[10] At this time Roth was investigating the tribes of northern Queensland and in 1902 he published the first illustrations of Australian figures. Unfortunately the procedures were not recorded, hence his large collection of seventy-four patterns is of little ethnological significance except in so far as it demonstrates the great wealth of designs and breadth of interpretation within a given area.

In 1902 Rivers and A. C. Haddon brought out a paper on method of recording string figures, a system on which all subsequent descriptions have been based, although many writers have substituted details in terminology.

---

[6] Hall, p. 366; Gill, p. 65.

[7] Tylor, p. 26.

[8] Klutschat, p. 136.

[9] Boas, p. 229.

[10] Tenicheff, p. 153; Kroeber, p. 298.

The first attempt to treat the subject from a universal point of view was by Caroline Furness Jayne in 1906. Mrs. Jayne not only brought together the material available at that time but also extended considerably our knowledge of the distribution and importance of string figures among the various peoples of the world by personally recording a great number from the several native groups assembled at the 1904 World's Fair at St. Louis. Since her pioneering effort many collections with full descriptions and illustrations have been made in many regions but only Miss Kathleen Haddon has evinced much general interest in their worldwide ethnological significance and only Jenness for the Eskimo has contributed a comprehensive comparative study of an entire ethnic area, although a number of preliminary surveys have been made for Polynesia, each from the point of view of some particular group of islands.[11]

AUSTRALIAN STRING FIGURES

Although Australia was the first continent for which string figures were reported it has been the most neglected in so far as this subject is concerned.[12] For other parts of the world we now have a fair number of descriptive studies, but for Australia the interest has been lamentably weak. Following Roth's illustrations in 1902 no interest was shown until Miss Haddon visited the continent in 1914 and later published in 1918 the descriptions of six figures from South Australia, one from Victoria and seventeen from Cape York, North Queensland. She was followed in 1925–1926 by Stanley who collected over sixty figures in the Cairns region of Queensland. Although there have been excellent opportunities for investigation in other parts of the continent they have been entirely neglected by other field workers.

In recent years a growing interest has been manifest but unfortunately it has been confined to the collection of final patterns rather than the recording of the procedures. However, these collections of finished figures do demonstrate the widespread importance of string figures in Australian culture and indicate the fruitful field which awaits investigation. Mounted on paper in the South Australian Museum is a collection by Mr. N. B. Tindale consisting of five South Australian figures, of which three are from the Coorong, two from the Diamantina district, and seventy-eight figures from Central Australia.[13] Of the latter at least sixteen are duplicates, although they are known by different names. Since these particular patterns are quite simple their procedures may or may not be the same, for similar results often can be obtained by different manipulations. Unfortunately the strings have slipped for some of the other figures, hence they are unreliable for comparative purposes.

---

[11] See Handy, Hornell, Dickey.

[12] Spencer, II, p. 581, who investigated the central regions where string figures have been found to be particularly abundant, makes reference to them only in passing for the Borraloola district, North Australia, and illustrates two finished figures. Bates, p. 95, states that Prof. Radcliffe-Brown in 1910 was concerned with string figures near Sandstone, Western Australia.

The museum also has a collection of sixteen mounted patterns gathered in Arnhem Land, North Australia, by Mr. H. U. Shepherdson. Although these designs are basically similar to other North Australian figures, a number of them are characterized by peculiar twists in the strings which are unique and presumably of local development.

Another district in which interest recently has been aroused is southern Queensland where Mrs. C. Tennant-Kelly has photographed a number of patterns.

The string figures described in this study were collected by the author in 1930, 1931, 1938 and 1939 in various localities in North Australia, Western Australia, South Australia and Victoria. Altogether one hundred and twenty-four string figures and twenty-one string tricks were recorded but as the result of duplications in various areas the total number of different examples is somewhat less. Their distribution among the several tribes and localities is indicated in the following table. Catches and cat's cradles were not noticed anywhere.

NORTH AUSTRALIA

	String Figures	*String Tricks*
Katherine River-Daly River-Victoria River District:		
Wardaman Tribe	41	2
Ngainman Tribe	13	
Wagoman Tribe	4	
Tagoman Tribe	1	1
Marithiel Tribe (Brinken)		1
Mullik-Mullik Tribe		1

---

[13] For a list of these figures see appendix. Mr. N. B. Tindale has informed the writer that string figures are made along the western coast of South Australia.

Darwin District (collected in Darwin):		
Melville Island	8	
Darwin		1
Coast, East of Darwin	2	
WESTERN AUSTRALIA		
Kimberley District:		
Djaru Tribe	7	
Kidja Tribe	7	
Broome-DeGrey District:		
Karadjeri Tribe	1	
Nangumarda Tribe	3	
Ngarla Tribe	1	
Targudi Tribe	1	
Northwest District:		
Panjima Tribe	2	
Injibandi Tribe	3	
Ngaluma Tribe	4	2
Mardudhunera Tribe	2	
Talainji Tribe	2	
Baiong Tribe	2	
Ingarda Tribe	3	6
Southern District:		
Kellerberrin	3	
Southern Cross	5	3
Quairading	1	
Gnowangerup	2	
Laverton	1	
Mogumber		2
Moora	1	
CENTRAL AUSTRALIA		
Arunta Tribe (collected in South Australia)	1	
SOUTH AUSTRALIA		
Maraura Tribe	1	
Swan Reach	1	3
VICTORIA		
Western District (collected at Lake Tyers, Gippsland)	1	

The locations of the tribes and localities mentioned are plotted on the map in Fig. 1. Also included in the positive distribution are other areas for which native testimony confirms the presence of string figures but for which specific procedures and patterns were not known to the informants. For instance, there is no doubt that string figures are present throughout the Northwest, the southern Kimberley Division and adjacent North Australia, and at least the western portion of the great desert east of the DeGrey River. Presumably there is a contiguous distribution throughout the continent but the blank spaces on the map must await more reliable information.

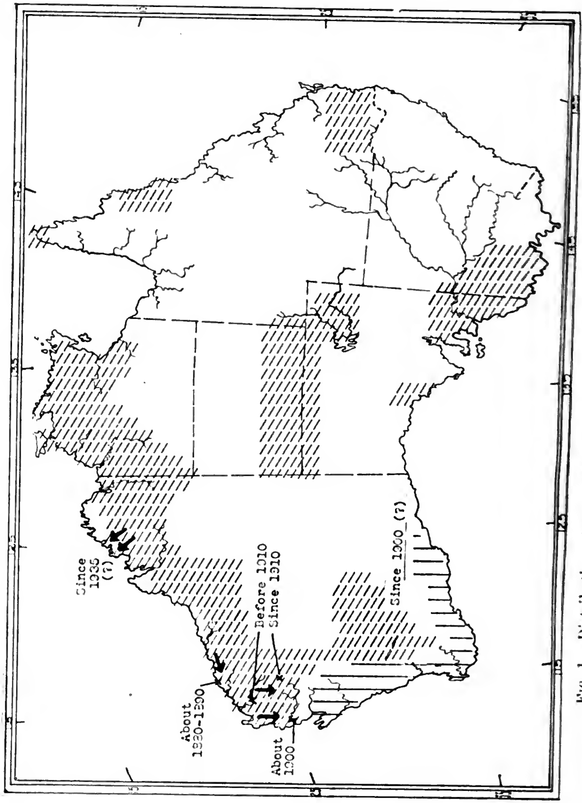


FIG. 1. Distribution of String Figures in Australia on the Basis of Information Available in 1939.

← Known directions of recent diffusion.

| | | | String figures known to be lacking. (Information from old natives, and from early white settlers at Busselton (age 85), Warriup, Esperance, Mingenew (age 84), Mullewa (age, over 70), Mount Magnet and Cue.

\\\\ String figures present. Tribes * and Areas include:

Queensland—
Cape York and Torres Strait.
Cairns District.
Southeast and Brisbane.

New South Wales—
No specific mention.

Victoria—
Central and Western Districts.

South Australia—
Northeast—Diamantina.
Southeast—Maraura and Adelaide tribes, Swan Reach and Coorong.
West Coast District.
Northern District—Arunta tribe and neighbors.

Central Australia—
Arunta tribe and neighbors.

Western Australia—
Kimberley Division—Kidja and Djaru tribes and neighbors.
Broome-DeGrey District—Karadjeri, Nangumarda and Ngarla tribes and neighbors.
Northwest Division—Ngaluma, Kariera, Injibandi, Panjima, Targudi, Mardudhunera, Noala, Jiwali, Warienga, Wajeri, Talainji, Baiong and Ingarda tribes and neighbors.
Inland Central and Southern Districts—Moora, Laverton, Kellerberrin, Southern Cross, Gnowangerup, Mogumber and Quairading.

North Australia—
Northeast—Arnhem Land and Borraloola.
North—Melville Island, Darwin and Coast.
West—Wardaman, Ngainman, Tagoman, Wagoman, Mullik-mullik and Marithiel tribes and neighbors.

* For more specific locations see Davidson, 1938.

The above list should not be considered an index to the total number of figures known in any particular tribe or locality but merely as a record of the number collected in each by the author. Knowledge of and interest in string figures, as well as ability to portray them, were found to vary considerably among individuals, hence to secure a complete list for any particular tribe a canvass of the entire population would be necessary. For the Arunta of Central Australia and those North Australian tribes represented in the table by only a few figures the data were gathered from stray individuals encountered in other tribal territories. Since it appears likely that all tribes in a particular district are acquainted with more or less the same figures it would seem that a more intensive survey would reveal that the North Australian Wagoman, Tagoman, Marithiel and Mullik-Mullik tribes, represented on the list by so few figures, actually know a great many.

The South Australian and Victorian figures were secured from among the few survivors of their particular tribes and it probably is now too late to supplement them by further inquiries.

In so far as Western Australia is concerned the above list in general probably is fairly accurate. The author had the opportunity to make rather intensive investigations in this state and although most of the tribes enumerated are now quite decimated and Europeanized, the figures obtained throughout a tremendous region were found to be quite consistently the same, a condition which suggests that other figures, if known, probably are not numerous.

As the result of the widespread distribution of many figures there are several duplications in the above tabulation. If these be eliminated by states the number of different figures in each, as represented in the author's collection, would be as follows:

	String Figures	*String Tricks*
North Australia	53	5
Western Australia	14	6
Central Australia	1	
South Australia	2	3
Victoria	1	

Further revision for the continent at large reduces the total collected by the author to sixty-eight string figures and six string tricks.

ANTIQUITY OF STRING FIGURES IN AUSTRALIA

That string figures were enjoyed in much of Australia in aboriginal times is certain. Their early presence is attested for various eastern and southeastern localities by the documentary evidence already mentioned and can be inferred for the central and northeastern portions of the continent from the great intensity of appearance in those regions at the present time. On the other hand, it is important to note that they are not of continental distribution. They are not reported for Tasmania, for which there are theoretical grounds for believing they were entirely unknown in aboriginal times, and it can be said definitely that they were lacking in the coastal districts of Western Australia between Esperance on the south and Shark's Bay on the north, as well as inland through the Murchison area apparently as far east as Mount Magnet and Cue. For most of the remainder of Western Australia, which occupies roughly one third of the continent, we have evidence to show that string figures are of recent introduction from areas to the east.

In the extreme northwest the relatively recent westward movement is well affirmed by native testimony. Old Ingarda natives in the Carnarvon area claim they first learned string figures in their childhood, about the beginning of the twentieth century, from their northern neighbors, the Baiong. The latter, in turn, state that string figures were unknown to them until about that time when they acquired them from their neighbors to the north, the Talainji. The general accuracy of this testimony is substantiated by old white inhabitants who recall the marked enthusiasm, about 1900 or shortly thereafter, which accompanied the initial introduction.

For the coastal area further to the northeast an old Mardudhunera man, apparently between sixty and seventy years of age, claimed that when he was a child, presumably some time between 1880 and 1890, his people first learned string figures from their eastern neighbors, the Ngaluma, who already possessed them at that time, according to the recollections of informants of about the same age from this tribe.

For the central Lyons River area, a few hundred miles to the south, an even more recent diffusion is indicated. Warienga natives of not over middle age assert that string figures came through the ranges from the north within their adult memory, or well within the present century. Among their northern neighbors are the Jiwali for whom a middle-aged representative, possibly in his early forties, stated that string figures antedated his memory. The claim of recent southward diffusion in this region is supported by an early white settler near Mount Augustus who insists that string figures were still unknown there in 1910, the year he departed the locality. Although it is generally found that whites who have employed natives for years and profess to know them intimately are unaware of the presence of string figures, a condition probably explained by the association of this pastime with leisure moments in the native camp rather than with working hours, most categorical denials of this sort cannot be relied on with any degree of certainty. In this instance, however, the statement is consistent with other evidence from the same general region.

The westward spread of string figures in the Northwest also is suggested by cultural evidence. The Ingarda and Baiong lacked nets in aboriginal times but both they and the Talainji, their northern neighbors and the westernmost tribe to make nets, apply the term "Net" to the same pattern. This figure is known by the same name among the net-making coastal Ngaluma and the neighboring inland Panjima and Injibandi of the central Fortesque River who lack nets but associate them with the Ngaluma. A similar situation in respect to the very same figure is found a few hundred miles to the east at Nullagine where the Nangumarda and Targudi who now reside there associate nets with the inhabitants of the lower DeGrey River.

A relatively recent introduction of string figures in most of Western Australia also can be inferred from the general tendency to interpret the patterns in terms of suggested European objects. Thus the figure so widely known as "Net" is termed by other inland tribes "Bed" because of its resemblance to the springs. For the same reason other designs in other peripheral districts are known by the same term. In certain southern localities "Gate" is a popular designation for other patterns. This type of evidence in itself is insufficient to warrant an interpretation of recency in those areas where string figures are numerous, such as North Australia where only a few designs have English names, for there probably is a common tendency to rename old figures as the aborigines become familiar with

European objects. However, since the use of English terms appears to be most pronounced in western peripheral districts for which we have other indications of recent introduction, this type of evidence is consistent with the conclusions advanced.

In the inland portion of southern Western Australia the antiquity of string figures is not definitely determined by direct evidence, but it would seem that a very recent introduction can be safely inferred. In all native camps in this area, as well as on various farms and sheep stations, there is a predominance of half-castes, most of whom as shepherds or itinerant laborers have travelled extensively in the state, often to northern districts where string figures are more common. The relatively few full bloods in this area, who also have moved about considerably, are now almost entirely detribalized. In addition one generally finds domiciled with them stray individuals who hail from all over the state, including the distant Kimberley district. Thus the ethnic situation at best is most confused and few informants can be considered reliable as to the exact provenience of the little information they are able to provide. Of the old people with whom the question was discussed all denied any knowledge of string figures in their childhood. The little positive information secured came almost entirely from individuals of middle age, a condition which suggests that their generation may have been the first to become acquainted with string figures. However, all were uncertain as to where or when they had learned the figures they performed. Some had the opinion they were white man's games. In recent times there seems to have been very little interest in them, for many informants stated they had not made them for years, nor had they taught them to their children.

One therefore gets the impression that the very few figures in these southern camps were not integrated with local aboriginal culture but belong to the more recent cultural melange of decadent local traits, stray aboriginal influences from other districts and predominant European civilization. Whether the local string figures have come from the northeastern part of the state or have infiltrated from Central Australia is not clear. However, it seems safe to assume that in these particular southern localities they are of quite recent appearance and probably arrived well within the present century.

Another district in Western Australia where string figure seem to be a recent arrival is the western Kimberley coasta country. Various white settlers claim that a few figures hav been introduced from south of the range only during the la few years. A general northwestward diffusion in this region i consistent with the traditions of the Djaru and Kidja tribes wh live south of the range, and who thus may be partly responsib for passing on some figures, for they associate string figures i general with the tribes to their east and southeast along th North Australian border. Incidentally it would seem that th Djaru and Kidja have not been acquainted with this pastim for long, for they know only a few figures of fairly commo appearance.

For Western Australia we therefore have reason to believe that string figures were entirely unknown in most of the stat until near the end of the nineteenth century. How much earlier they may have come to the districts directly contiguous to North Australia and Central Australia cannot be indicated for lac of data.

ANTIQUITY IN EASTERN AUSTRALIA

The antiquity of string figures elsewhere in Australia is not demonstrated by direct evidence, but several consideration point to northern Queensland as the area whence they diffused to other parts of the continent.

For instance, there seems to be a definite tendency for the number of figures to decrease as the distance from northern Queensland increases. For the latter area between one hundred and one hundred twenty-five already have been reported, compared with sixty-eight for North Australia, eight for the Kimberley district and a total of six for the remainder of the Northwest. In a southwesterly direction we find a minimum of sixty-three for Central Australia, seven for southern Western Australia. For southern Queensland the number has not been determined and no data are available for New South Wales, but for eastern and southeastern South Australia eleven patterns have been noted, for Victoria two.

It seems quite certain that these data greatly understate the situation for all areas, except possibly Western Australia, for it cannot be doubted that the total number of figures in each is or previously was somewhat if not considerably greater. However, since this condition also applies to northern Queensland where the large number of known figures come from a few small localities, the present evidence, unsatisfactory as it is for most areas, may represent a fair sampling of relative intensity for the continent at large.

Another consideration which suggests northeastern Australia as the region of greatest antiquity is the great number of local patterns there. In most areas where string figures are numerous, a situation which suggests pronounced popular interest, a relatively large proportion seem to be of quite local appearance. All factors being equal it can be assumed that such a condition is indicative of a longer experience with string figures than is the case where only widespread patterns are found. In this respect it seems significant that in Western Australia no distinctly local designs have been identified, nor are there any peculiarities in manipulation not typical of other parts of the continent. On the other hand, among the figures of northern Queensland, North Australia and Central Australia are many patterns not noticed as yet in other regions. Thus it seems that experimental interest has led to successful results in some Australian districts. If there be readers who consider the development of new patterns a simple achievement, to say nothing of the discovery of new steps in manipulation, let them try their hand to this end.

It should not be inferred that attainment of new figures necessarily involves radical departures from standard Australian openings, movements or extensions. There are some novel elaborations, such as the remarkable twists which characterize several figures from Arnhem Land, North Australia, and presumably others are awaiting recognition, but judging from the basic character of the patterns and procedures now known, it would seem that a large percentage of new Australian designs has been developed by varying or combining the procedures of older figures. Thus, an added or variant step at the beginning of manipulation may cause quite a different final pattern, but one which nevertheless is fundamentally of similar character to other Australian figures. At the moment it is difficult to analyze and define satisfactorily these distinctive characteristics, for we are in need of more comparative material from other parts of the world before it will be possible to distinguish regional peculiarities. But if one will acquaint himself with the Australian figures in this monograph and those from some other part of the world, for instance the Eskimo collection described by Jenness, he will sense many points of difference not only in the more obvious manipulations but also in the contrasting types of interlocking arrangements in the patterns. If put to the test of identifying string figures of unknown provenience it seems not unlikely that a specialist would be able on technical grounds to reject many figures as not belonging to his area and to accept others as peculiar to the major ethnic region of his acquaintance. Such distinctions obviously would have to be made in terms of very specific secondary criteria for in all parts of the world the basic string figure traits are more or less identical.

Naming and Function of Figures

In Australia, as in most other parts of the world, the vast majority of string figures have been named after animals or objects. Such a selection would seem to be natural for a people who lead a primitive hunting existence. On the other hand, in view of the complex social and totemic systems in Australia and the importance of magic and food-increase rites in native life, it seems strange that practically no patterns have been given a social, totemic or magical connotation.

On the basis of present information it would seem that string figures in Australia serve almost entirely the purpose of amusement of young and old of both sexes, with the women often knowing a greater number of figures than the men. The only reported exception are the figures of Arnhem Land, North Australia, concerning which there is a vague rumor of some association with initiation ceremonies.

String Tricks

There is little to be said in respect to the six string tricks present in Australia. All are of worldwide appearance and their widespread distribution in Australia suggests that they are not the result of European introduction. String catches were not encountered by the author on the continent and have ot been reported as yet by other investigators. Examples have een collected in Torres Strait and Oceania, hence their disovery on the continent would come as no surprise.

MELANESIAN DERIVATION OF AUSTRALIAN STRING FIGURES

The principal reason for considering northern Queensland he area of greatest antiquity is found in a comparison of Australian and Oceanic figures, for on this basis there can be no doubt that string figures first came to this part of the continent from Melanesia. Not only do we find in Melanesia identical procedures for a number of the more complex Australian patterns but, equally significant, also the presence of distinctive string figure usages, manipulations and extensions not reported for other parts of the world. As stated above it is difficult to define at the moment many of these peculiarities but the evidence as a whole seems sufficient to indicate that Australia, Melanesia, Micronesia and Polynesia comprise a major string figure area.

Since many designs can be produced by entirely different procedures, comparisons of final patterns, unless accompanied by full descriptions of the manipulations, are worthless for historical purposes.[14] For this reason Roth's seventy-four Queensland designs must be excluded from consideration. This leaves North Australia as the area most abundantly represented, hence most of our comparisons concern figures from this region. We have noted that both Queensland and North Australia are characterized by a large proportion of local figures, nevertheless there are several patterns, Nos. III, X, XXXIV, XXXV and LXVIII, known to be constructed by the same manipulation. When more extensive collections are available it seems likely that the number of figures with identical procedures in the two areas will be increased.

Of the figures fully described for the continent the following are made in the Oceanic areas indicated by identical manipulations, except for very minor variations in a few examples:

---

[14] For example, one of the most widespread figures in Australia, Melanesia, Micronesia and Polynesia, No. XXXV, is produced by identical procedures throughout. The final pattern, however, is found also in North America and Africa where the Klamath and Yoruba respectively construct it by entirely different manipulations.

Nos. IX, XV and XXXV in the D'Entrecasteaux Islands.
Nos. IX, XXXV and XXXVIII in New Caledonia.
No. XXXVIII in New Guinea.
Nos. III, IX, XV, XX, XXXII, XXXV, XXXVI, XXXVIII (similar) and LXVIII (similar) in Yap.

Some of these figures also appear in Polynesia but the problem of their presence there is associated more with Melanesia than with Australia and need not concern us at the moment.

The finding of a greater number of identical procedures in Yap, over 1,600 miles north of Australia, than in the nearer Melanesian islands, is interesting and possibly may be an important clue to the history of string figures in the whole Oceanic region. This does not mean that there is any direct association between Yap and North Australia for it may be merely coincidental that these particular patterns were among the first collected in each area. It seems likely that all are known in the intervening area, for several already are reported there as a glance at the above list indicates.

Less obvious than the procedures of specific patterns for showing historical relationship, but perhaps of greater ultimate value for this purpose, are a number of string figure practices and peculiarities found in one part of the world but not in another. Since string figures and the characteristics of their construction are unknown or very poorly known for so many parts of the world any attempt at present to define technical traits and plot their distributions can be only of a most preliminary nature. The following traits definitely indicate uniformity in tradition between Australia and Oceania, but whether they are peculiar only to these two regions, either specifically as stated or as redefinable Pacific variants of more widespread customs, cannot be determined at the moment. This list, therefore, is offered tentatively, subject to revision and redefinition:

1. Use of two strings and the services of two players for a single figure. See LXVIII. North Australia, Queensland, Yap, New Zealand, Marquesas and Hawaii. Compare with final patterns from Palau and New Guinea. (For references see Appendix E.)

2. Placing strings on lap for certain movements. See XLI, XLIII and XLVI. North Australia, D'Entrecasteaux Is. and New Caledonia.

3. Crossing hands for extension. See III. North Australia, Queensland, Torres Strait, Fiji and Yap.

4. Turning one hand in, the other out, in some extensions. See II, XXIV, XXXIX and XL. Much of Australia and Oceania.

5. Use of doubled loop. See I, II, III, VIII. North Ausralia and Yap.

6. Concept of a series whereby several progressive patterns are made by a single player or alternately by two players. See XLVII–LI and LII–LX. Australia, Fiji and the Gilbert Islands.

7. The Elbow Opening. See LXV–LXVII. North Australia and Palau.

8. The Pindiki movement.

A most important trait in Oceania and Australia is the Pindiki movement,[15] otherwise known as the Caroline Islands Extension. Since it is employed as a step in manipulation of many figures as well as an extension, the former term is preferred. A description of the movement is given on page 797. Our interest in it at this time is confined to its significance in Oceanic and Australian string figure construction.

The distribution of Pindiki is interesting and ultimately may throw important light on the history of string figure elaboration in the central Pacific. In addition to its apparent universal use in Australia it is found in New Caledonia, the Loyalty Islands, Fiji, New Zealand (not common), the Ellice, Tokelau, Gilbert and Caroline Islands, and in Nauru. It seems to be lacking in Tonga, the Marquesas and Society Islands, Tahiti, Hawaii, and in North and South America, Asia, Africa and Europe.

Whatever may be the antiquity of string figures in the Pacific the regional distribution of Pindiki suggests that this movement is of relatively recent origin. Since it is lacking in marginal Polynesia it may be of recent appearance in western Polynesia but for the time being we cannot determine whether it originated there or diffused from Melanesia or Micronesia. Since Pindiki is of universal use in Australia it seems likely that it accompanied the original introduction of string figures from Melanesia. If so, any conclusions which may be developed in respect to the time and place of the origin of Pindiki would throw light upon the question of the antiquity of string figures on the continent.

---

[15] From native terminology. See Haddon, K., 1930, p. 156.

The Pindiki movement, like other culture traits, appears to have undergone its own historical spread without any necessary association with any particular figure or group of figures. To a person learning string figures directly from natives this movement may appear as an integral part of the procedure. This association, however, is purely incidental, for the design could have been made by clumsier manipulations. In reality, Pindiki is no more than an efficient and convenient method of stretching the strings the better to portray the design. It can be applied to numerous figures in all parts of the world and for this reason its localized distribution cannot be ascribed to any peculiarity of the string figures of Oceania. Why this simple device of portrayal has not been discovered in other parts of the world is indeed remarkable.

BASIS FOR COMPARATIVE STUDIES OF STRING FIGURES

To those not intimately acquainted with string figures it may seem strange that the similarities between Melanesia and Australia are not more numerous. However, it should be pointed out that our collections not only are incomplete but the large proportion of local patterns in each area provide no basis for comparison. Nor has mention been made of the great number of identical fundamental string figure traits in Melanesia and Australia which, as the result of their universal appearance, throw no special light on Oceanic problems.

From the point of view of worldwide comparisons the identical appearances in Melanesia and Australia become much more impressive, for in spite of the fact that string figures in all parts of the world have a common foundation we seldom find the same patterns in widely separated regions, and of the few so found most are produced by entirely different procedures. Theoretically this condition is most unexpected for we should anticipate some evidence of a common substratum in simple procedures and patterns, as well as some examples of independent parallel development as the result of the laws of chance, for in all areas there are a number of constant factors in the principles of string figure construction which favor the accidental attainment of the same final design by the same means. These include:

1. *Openings.*—The construction of all string figures commences with a specific opening or initial arrangement of the strings on the fingers. As anyone can soon discover, it is possible to place a loop on the hands in many different ways. However, string figure makers throughout the world have utilized hardly more than a score of different arrangements and of these several represent elaborations of four fundamental types herein called First Position and Openings A, B and C. Indeed, since Opening A is derived from First Position, it could be maintained that there really are but three. Furthermore, it is theoretically possible that Opening B has been derived from Opening A. First Position and Opening A seem to be everywhere, as is Opening B. Opening C, which technically could be related to Opening B, occupies almost as widespread a distribution. The few odd openings appear only in regional or local distributions, are always found in association with the fundamental openings, and form the basis for so few figures that they are of little relative importance. Thus practically all the string figures in every region have been developed from four standard openings.

2. *Size of Loop.*—Another basic trait of universal uniformity is the size of the loop employed. In all regions most loops are made from a string varying in length between four and seven feet, but selected at random without any particular attention to exact size. Depending on the opening selected such a loop places the hands between eight and eighteen inches apart when manipulation commences. As construction proceeds the network of strings between the hands becomes more complex and as a result the hands are drawn closer and closer together. Therefore, other factors being equal, the smaller the loop the fewer the possible number of manipulations before further movements become impracticable, contrariwise the larger the loop the more numerous the possible manipulations and the more intricate the final patterns.

Since we have in the figures of many areas abundant evidence to indicate interest in the development of more complex designs, it seems strange that so few aboriginal experimenters recognized in a larger loop the answer to some of the problems with which they were faced, for only occasionally are we informed that a loop somewhat larger than usual is required or desired for a certain figure. Only the natives of Nauru, an Oceanic island seem to have seized upon the opportunities offered by a loop twice normal size and it is among their figures that we find a most extraordinary intricacy of pattern.[16] Whatever the abilities of these islanders in discovering involved procedures, there can be no doubt that their successful application can be attributed to the increased size of the loop they employed. If such remarkable results can be attained by the use of a loop twice normal size what labyrinths in string could be created if the possibilities of a loop several times as large were fully exploited!

There seems to be no rational explanation for the universal use of the same sized loop. It cannot be admitted that the customary size is employed solely because it is most convenient to handle, for were it not suitable for the vast majority of figures made by each people its length presumably would soon be altered. We should expect to find string figure dilettantes in each society content to accept the traditional loop without thought of the possibilities of more complex constructions, but it does seem peculiar that the experimenters responsible for the development of complex figures failed to recognize and take advantage of a great opportunity for much more intricate accomplishments.

3. *Principles of Constructions.*—With very few exceptions [17] final string figure patterns are suspended on two main transverse strings, usually the two parallel hand-to-hand strings found in each opening arrangement, but occasionally the two diagonal strings.[18] Hence, regardless of how these particular strings may be manipulated in the course of construction, they must be brought to their proper relative positions before the final figure can be completed. In the more simple figures this may be accomplished by only two or three movements; in the more complex examples many intermediate steps may be required.

An important consideration in construction is the coordination or lack of coordination of the hands. Since it is much easier to follow procedures in which the movements of the hands are the same, it would seem that those in which the manipulations of one hand are radically different from those of the other represent a relatively late development in string figure history. This conclusion undoubtedly is valid in respect to such an intricate figure as "The Fox and the Whale" of the King Island Eskimo, produced by twenty-eight steps in which the movements of the hands are decidedly unalike.[19] Such a complicated procedure can be the outcome of only a long experience with string figures and presumably requires an intelligence better than average for successful completion. Examples of comparable complexity have not been noted as yet for Australia, but a Torres Strait figure, "Trigger Fish," can be classified in this advanced category.[20]

---

[16] For illustrations, see Jayne, pp. 367–370. Barton, Pl. XXVI, indicates that a long string is used in British New Guinea for figures described as "perplexingly complicated."

[17] For example, see Figs. 9, 25–30, 40–42, 51, 99.

[18] The two diagonal strings are sometimes brought to a parallel position by a final crossing of the hands. For example see Nos. III and XXXIX.

Not all figures in which the movements of the hands are not coordinated are characterized by complexity. A much greater number are attended by relative simplicity and the differences in the movements may be confined to only one or two steps, often a manoeuver to pick up a particular diagonal string which occupies a different position on each hand. Obviously such procedures cannot be considered advanced on the basis of complexity of construction. On the other hand they certainly cannot be regarded as preliminary, for it seems reasonable to suspect that some general knowledge of string figure construction underlies the discovery of these special types of uncoordinated manipulations.

A third variety in which the movements of the hands are dissimilar is to be found in figures constructed almost entirely by manipulation with one hand.[21] In most examples the procedure is not extremely complex, hence here again there are no a priori technical grounds for considering the construction as particularly primitive or advanced. However, it is important to note that the figures of this variety are so few, yet so varied, that we are warranted in suspecting that they are not fundamenal in string figure history but more likely represent one or more tangent developments. Distributional evidence ultimately may provide the most important clues to this problem.

---

[19] Gordon, p. 94.
[20] Jayne, p. 96.
[21] For example, see Figs. 17, 18, 26, 27, 40, 42.

The question of coordination is important in its influence on the style of final pattern, for since the standard openings are characterized by a complementary arrangement to the right and left of the middle of the figure, coordinated movements are bound to produce a symmetry in final pattern. On the other hand, procedures in which the manipulations by the hands are dissimilar may lead either to an unbalanced design or, if one set of uncoordinated movements is countered by another which restores the original relative position of the strings involved, to a symmetrical pattern.

Four types of symmetrical pattern can be recognized:

Type 1: With right and left halves symmetrical in respect to a vertical axis.[22]

Type 2: With four quarters symmetrical in respect to both vertical and horizontal axes.[23]

Type 3: With ulnar and radial halves symmetrical in respect to a horizontal axis.[24]

Type 4: With two right triangles which coincide if one is turned 180°.[25]

Theoretically we should expect to find Type 1 the most common, as it seems to be, for such designs are the natural results of procedures in which the movements of the hands are coordinated.

Figures of Type 2, with a few simple exceptions,[26] require a more complicated construction for, in addition to the development of a lateral balance as in Type 1, a vertical balance must be attained by countering each move of an ulnar string by a reversed move of the complementary radial string.

Type 3 is not represented by many examples and of the few known most seem to be either figures extended on what originally were the diagonal strings or those constructed almost entirely by manipulation with one hand.

Figures of Type 4 seem to be quite rare. The procedures of some are fairly simple, of others quite complex.

---

[22] For example, see Figs. 8, 9, 13, 55, 57, 70, 73, 80, 81, 84, 85, 87, 88, 90, 92 and 96. See also footnote 26.

[23] See Figs. 31, 33, 37–39, 56, 59, 60, 83, 91, 95, 97. See also footnote 26.

[24] See Figs. 17–19, 25–29, 42, 43, 64, 66.

[25] See Figs. 14 and 22.

[26] There are several patterns which by definition should be classified in Type 2 but which because of their simplicity of construction are more like Type 1. See Figs. 44, 61–63, 69, 72, 86 and 89.

FIGURES OF ACTION

Although most string figures assume a fixed character upon completion, there are a few in which action is illustrated. There seem to be at least two varieties, the one in which action leads to a definite termination of the figure, the other in which the action can continue indefinitely. The first is found in Australia in Figs. 18 and 19, the second in Fig. 41.[27] Action figures have been noted in several parts of the world but all seem to be quite different in detail.

LACK OF UNIVERSAL UNIFORMITY IN PROCEDURES AND PATTERNS

In spite of the fact that virtually all string figures in all regions have been developed from a universal substratum of three constant factors it is indeed remarkable to find so little similarity in the several hundreds of procedures now recorded. The numerous differences in patterns developed from the same opening obviously are attributable to two variable factors, (1) the types of manipulation, and (2) the number of manipulations. The number of manipulations, as already noted, can be limited by the size of the loop, but since most recorded figures have been developed from a loop of standard size, this consideration is of more theoretical than practical importance. The types of manipulation have not yet been classified but generally they are concerned primarily with the diagonal strings which ultimately become transformed into a central figure suspended on the two original parallel strings.

Unlike most culture traits, which generally serve a number of purposes, the three constant and two variable factors in string figure construction apparently have no other importance than as constituent elements for a final pattern. Individually they have no significance when removed from their special context or if taken outside the area of string figure construction. Since most string figures cannot be made without the combination of these various elements, it is clear that they comprise a complex which must travel as a unit whenever string figures invade a new region. On the other hand, it also is apparent that these factors can exist individually and diffuse individually within the distribution of string figures. Thus a new opening, an innovation in manipulation, a dextrous or simplified movement, a different method of extension, and the like, can be detached from the particular context in which they originated and diffuse individually and without reference to other traits within the distribution of string figures in general. Presumably the spread of such an element often, if not usually, would accompany the diffusion of a particular procedure, but such an association is not a requisite. For instance, the convenient Pindiki movement could easily be taken over and successfully applied to old procedures. Similarly the diffusion of a new procedure in most cases would be no more than the diffusion of a different way of combining traits already possessed by the borrowing group.

---

[27] Roth, Pl. 3, Fig. 5, gives an example of the first in "Man Climbing a Tree." Haddon, K., 1918, p. 129, describes one of each in "Making Fire" and "A Corroboree."

When our collections of string figures become more numerous it seems not unlikely that some fairly complex examples of identical procedure may come to light in widely separated regions. As already indicated such appearances should be not unexpected in view of the possibility that old procedures of common origin may still be retained in various regions, as well as from the possibility that technical considerations favor the chance development of parallel procedures in two or more areas. However, the problem of determining which may be responsible in any given instance is indeed formidable and will require more carefully selected criteria than can be defined at the present time. At the moment it would seem that the only reliable indication of historical relationship is actual or inferable contiguity in distribution in a region characterized by similarity in secondary traits.

The remarkable diversity in string figure patterns in the world today is impossible to explain except in terms of a long period of time. As new figures are invented, possibly old figures become forgotten, hence in the course of centuries or millenia the originally similar appearances in two different areas may change so completely that the more obvious factors such as the procedures and patterns no longer show any resemblances. On the other hand, the more obscure technical traits may still reveal the clues to the basic question of derivations and it is in them that we may find our most valuable evidence for such problems.

The worldwide distribution of string figures suggests that their origin must be sought in considerable antiquity. Theoretically no maximum antiquity can be indicated, for since the only technical background for string figures is a knowledge of string, it is possible that they were made in Palaeolithic times from rawhide string or babiche. The situation in Asia is interesting for since the several non-contiguous appearances are found just beyond the boundaries of recorded history it is possible that some mention of the pastime may be located in old Chinese records. For the New World and Africa it seems quite probable that several thousands of years must be allowed for the diffusion from some Asiatic point of origin to the present appearances. On the other hand, the situation in Oceania does not necessarily presuppose a great antiquity. The lack of string figures in Tasmania and in coastal Western Australia and the apparently rapid diffusion through the western third of the continent within the last century, suggest that they could have been introduced into the continent only a few hundred years ago. Thus it is possible, but by no means demonstrated nor necessarily probable, that string figures came into the Pacific with Polynesians or Micronesians, or diffused from the west into New Guinea and western Melanesia late in the pre-Polynesian period. Distributional evidence ultimately may reveal some definite suggestions on this question. By comparison of secondary traits it may be possible to indicate an approximate antiquity for string figures in Australia.

NOMENCLATURE

The system of description and nomenclature employed herein follows in a general way that used by Mrs. Jayne with modifications taken from Roth (1924). This terminology is less technical than that recommended by Rivers and Haddon and should be less confusing to the lay reader. The various terms with which one should be familiar include "Usual position," the several "Openings," "Navaho," "Pindiki," "Extension" and "Normal release." In all descriptions the positions of the strings are given from the point of view of the manipulator and not from that of the spectator who usually faces the former and to whom, therefore, the strings appear in reverse order.

Usual Position

For "usual position" the manipulator holds his hands in front of him with palms facing, fingers outward. Most figures commence from this position and as a general rule it is returned to after the completion of each step in the procedure.

THE STRINGS

In the descriptions the term "string" is applied to a single cord wherever it may be found. Thus one may have a "string" on the near (radial) side of the thumb or a finger as well as a "string" on the far (ulnar) side of the same digit. Such strings are spoken of as the near thumb string, the near index finger string, the far little finger string, and so on. Two strings on the same digit, if they are the continuation of one another, form a loop and may be referred to as the thumb loop, middle finger loop, etc. When there are two strings or two loops on the same digit, that which is nearer the tip of the finger is called the upper (distal) string or loop, that nearer the base of the finger, the lower (proximal) string or loop. A string across the palm is spoken of as a palmar string, that on the back of the hand, a wrist or dorsal string or loop.

OPENINGS

The "openings" or initial arrangements of the strings used by the Australians are the same as those usually found elsewhere in the world. These include "First position," "Opening A," "Opening B" and "Opening C." In addition, in restricted areas odd openings may be present.

FIRST POSITION

The "first position" is the most common of all openings throughout the world, for it not only is the arrangement from which many figures are developed directly but it also is basic to Opening A.

1. *Left hand.* Hold palm downward, fingers outward. Pick up the string on the back of the thumb so that a long loop hangs below.

2. *Right hand.* Hold palm downward, fingers outward. Move the hand away from you and insert thumb under the string which lies across left thumb.

3. *Both hands.* Move hands apart until the strings are taut. You now have a loop on each thumb. Insert little finger from

low into thumb loop, pick up on its back the far thumb string d return to position (Fig. 2).

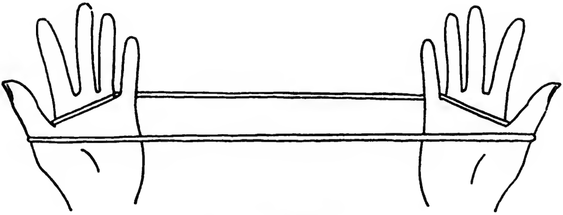


FIG. 2. First Position.

OPENING A

Opening A also is extremely common in all parts of the world. The steps in its construction are reminiscent of Cat's Cradle but the initial arrangement of the strings is not the same.

1. *Both hands.* First position.
2. *Right hand.* With back of index finger pick up from below the left palmar string and return to position.
3. *Left hand.* With back of index finger pick up from below the right palmar string between the strings of the right index finger loop. Draw out and return to position (Fig. 3).

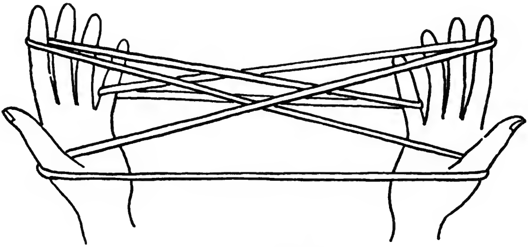


FIG. 3. Opening A.

OPENING B

Another common Opening found among most peoples who construct string figures is here called Opening B. It is also known as the "Top Opening."

1. *Both hands.* Hold string between tips of thumb and index finger with palms downward and hands about ten inches apart. A long loop will hang below.

2. *Right hand.* Move hand forward thence to left and back until tips of thumbs meet. Pick up between the tips of right thumb and index finger the string held by left thumb and index finger. You now have a small loop held by right thumb and index finger.

3. *Left hand.* Pass index finger over the strings held by right thumb and index finger, insert from the far side into the small loop and bring tip against tip of left thumb.

4. *Right hand.* Roll tips of thumb and index finger until nails meet to permit the strings to be encircled as on the left hand.

5. *Both hands.* Draw hands apart keeping what was the small loop on index fingers and what was the long loop on the thumbs. Separate thumb and index finger, move thumb upward and turn hands palms facing (Fig. 4).

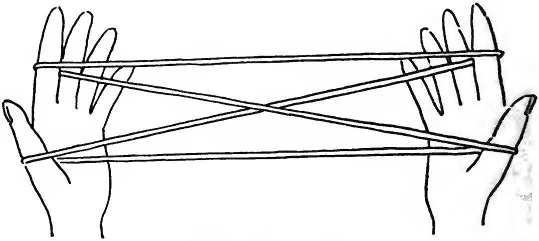


Fig. 4. Opening B.

Opening C

This opening, although not as common as Opening A or B, is also widely distributed in the world. It is known as "Little Fishes" Opening.

1. *Both hands.* Pick up string between tips of thumb, index and middle fingers, hands about ten inches apart. Make a loop in the string by bringing the hands together, the right hand a little forward of the left. Insert the index finger into this small loop from the far side until string is at base of finger. Press middle finger against index finger to prohibit string from

sliding up finger. Close ring and little fingers loosely over pendant string. Reach index finger toward you into long hanging loop and seize between tips of thumb and index finger the string which runs across hand to the far side of the little finger. Extend the hands, slowly playing out this string between the thumb and index finger and keeping the first loop picked up well at the base of the index finger. When the hands are fully extended keep the strings taut, remove all fingers except index finger from strings, rotate index finger inward and upward by lowering elbows and extend (Fig. 5).

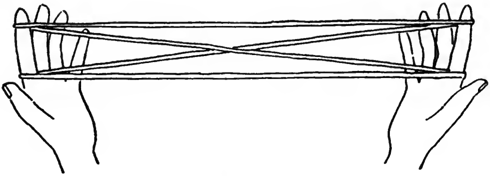


FIG. 5. Opening C.

ELBOW OPENING

This unusual opening has been noted only in western North Australia and Palau in Oceania.

1. *Right hand.* Hold string loosely on hand, palm upward, fingers to left, so that a long loop hangs below. Pass loop over left hand and forearm and suspend above left elbow (or variantly, on the shoulder). Withdraw right hand. (See Figs. 99–101.)

PINDIKI

In the construction on many Oceanic and Australian figures there is a distinctive movement called "Pindiki" as already mentioned. It also is known as "Caroline Islands Extension" but, since it is employed as a step in manipulation as well as an extension, the former term seems more satisfactory.

The most simple arrangement of the strings for Pindiki consists of two near thumb strings, one far thumb string, a palmar string and a far little finger string on each hand (see Number XLII, Movements 1 and 2, and Fig. 6). Additional strings may be present but the ones enumerated serve as the pattern for the manipulation. The movement is simple and is performed as follows:

Both hands. Hold hands in usual position. Bend index finger down over palmar string, straighten it toward you, and

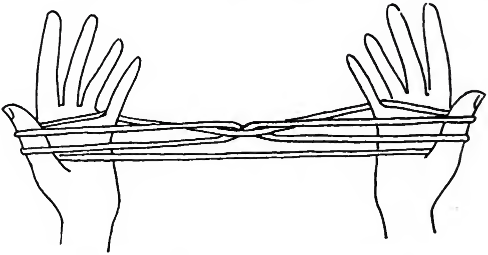


FIG. 6. A Step in Executing the Pindiki Movement.

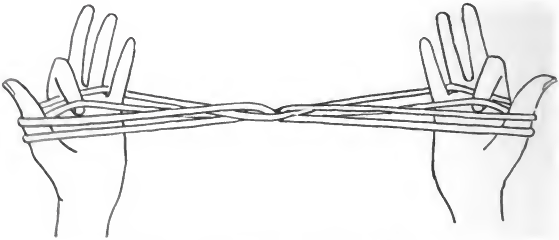


FIG. 7. A Step in Executing the Pindiki Movement.

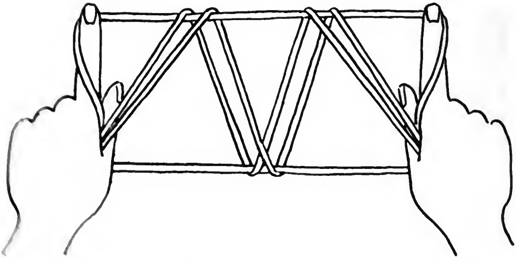


FIG. 8. The Pindiki Extension.

pick up from below on its back the far thumb string (Fig. 7). Hold this string secure against the base of the index finger by pressure from the thumb and extend the index finger to permit the string on its back to slip on to the back of the first knuckle. Turn the hands palms outward and extend by working the strings between the ball of the thumb and the base of the index finger (Fig. 8).

NAVAHO

The term "Navaho" is now well established in string figure literature in reference to a movement first recognized as a distinctive step among the Navaho Indians, but since found to be extremely common in most areas where string figures are made. The principle is simple and involves the lifting of a lower string over an upper string on the same digit and dropping it on the other side of the finger. Since this manipulation is performed by everyone in a given society its character is social. However, the manner of its performance varies considerably in terms of individual skill and dexterity. In some cases the manipulator employs his teeth or lips to lift up the lower string, in others this string is grasped between the tips of the thumb and index finger of the opposite hand while the other fingers are closed against the palm to secure the strings on that hand. The more able manipulators raise and lower the elbows and rotate the digit involved to cause the lower string to slip over the upper string.

RELEASE

The term "release" is applied to any movement by which a figure can be converted directly to the original loop without causing snarls in the strings. In a large proportion of string figures this result is accomplished by laying the figure carefully on the lap free of the fingers and drawing out at top and bottom the two main transverse strings at their middle points. In such cases the release is spoken of as "normal release."

STRING FIGURES FORMED FROM FIRST POSITION

I. TWO DOGS

WARDAMAN.

1. *Both hands.* Use doubled string or short loop. First position.

2. *Both hands.* Drop thumb strings. Pass thumb from above between little finger strings and pick up from below the far little finger string.

3. *Both hands.* Hold hands with strings slack. Bend thumb over palmar string and pick up from below the near little finger string close to base of little finger. Draw out over palmar string and by turning the palm downward and elevating the elbow allow the lower thumb loop to slip over the upper thumb loop; that is, Navaho the thumb loops. Turn hands palms facing and extend (Fig. 9).

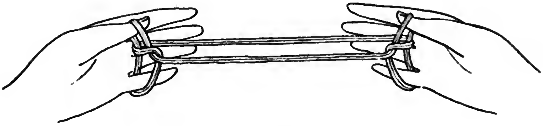


FIG. 9. Two Dogs.

Release normal

II. A WATER HOLE

WARDAMAN, NGAINMAN.

1. *Both hands.* Use doubled string or short loop. First position.

2. *Right hand.* Pick up from below with back of index finger the left palmar string and draw out. With the back of thumb pick up from below the near right index finger string.

3. *Left hand.* Pick up from below with back of thumb the left near little finger string. Bend index finger down over left palmar string and pick up from below the left lower far thumb string.

4. *Both hands.* Navaho thumb strings and drop strings held by little finger. Turn hands palms downward.

5. *Left hand.* Turn downward and inward until palm is facing inward with fingers pointing to right.

6. *Right hand.* Pass right hand upward to left by raising elbow until directly over left hand (Fig. 10).

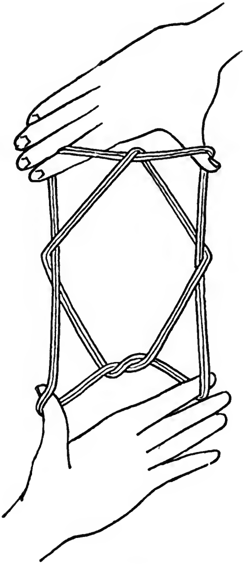


FIG. 10. A Water Hole.

III. TWO WATER HOLES

WARDAMAN, NGAINMAN.

This figure is identical in procedure with Queensland "Snake," Yap "Flint and Steel" and, except for minor variations, Fiji "Utosuvi." Comparison also should be made with Queensland "Deaf Adder," procedure not given, and a figure from Torres Strait (See Chart at end). In North Australia as in Yap a doubled string is employed.

1. *Both hands.* Use doubled string or short loop. First position.

2. *Right hand.* Pick up from above with index finger the left palmar string. Rotate the index finger toward you to give a half twist to the index finger loop. Extend.

3. *Right hand.* Pick up from below on back of thumb the near index finger string.

4. *Left hand.* Pick up from below with thumb the near little finger string. Bend index finger over left palmar string and pick up from below the far thumb string. Return to position.

5. *Both hands.* Navaho thumb strings. Drop little finger strings and turn hands palms downward.

6. *Left hand.* Rotate downward then upward to the right until palm is facing upward and back of left wrist touches back of right wrist.

7. *Right hand.* In conjunction with Movement 6 move right hand to the left and turn palm upward (Fig. 11).

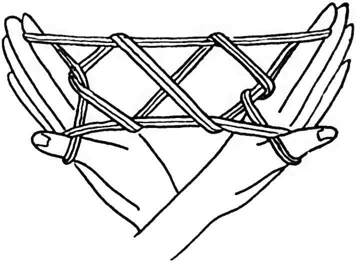


FIG. 11. Two Water Holes.

IV. A DEAD MAN IN A TREE

WAGOMAN.

A long string is required for this figure.

1. *Left hand.* Hold loop as in first position with hand in front of shoulder.

2. *Right hand.* With thumb and index finger pick up near left thumb string and draw under chin, then upward over top of head, downward under chin again. Repeat this encircling movement until only a short length of string remains. Pick up this remaining portion from below as in First position (Fig. 12).

3. *Both hands.* Hold hands as in First position and pick up with teeth the far little finger string. Hold strings securely

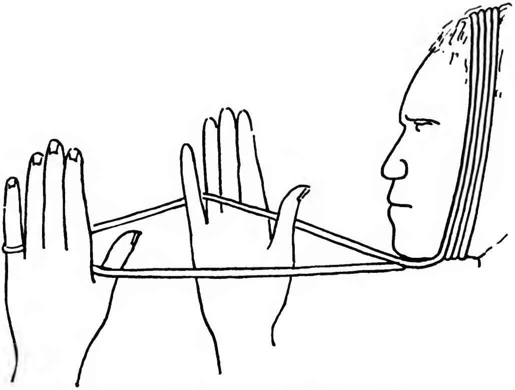


FIG. 12. A Step in "Dead Man in a Tree."

and reach each hand to the back of the ear and pull down the strings wrapped around the head so that they fall over the strings held in the mouth and between the hands. Drop thumb loop. Bend thumb back and pick up from below the loop held in the mouth. Release teeth and extend (Fig. 13).

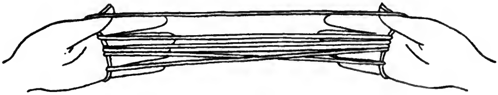


FIG. 13. Dead Man in a Tree.

V. PHALLUS

WARDAMAN.

1. *Both hands.* First position.

2. *Both hands.* Drop thumb strings. Reach thumb under the little finger strings and return with both on back of thumb.

3. *Right hand.* With index finger pick up from below the two left palmar strings and draw out.

4. *Left hand.* With index finger pick up from below the two right palmar strings and extend.

5. *Both hands.* Drop thumb strings and rotate index finger outward several times to twist index finger strings.

6. *Right hand.* Pass hand downward with palm toward body. Extend fingers.

7. *Left hand.* Pass hand inward and upward with palm facing outward and extend fingers (Fig. 14).

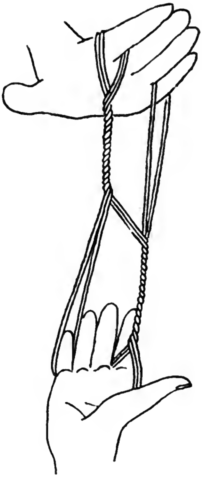


FIG. 14. Phallus.

VI. DILLY BAG

WARDAMAN, NGAINMAN.

Compare final pattern with Queensland "Dilly Bag" and Arnhem Land "Mutchichi," the procedures of which are not given. In Arnhem Land a doubled loop is employed.

1. *Both hands.* First position.

2. *Right hand.* Place loop over head and withdraw hand. Bend index finger over left palmar string. Draw out and place in mouth. Withdraw index finger (Fig. 15).

3. *Left hand.* Bend thumb over far thumb string and pick up from below on its back the near little finger string. Extend

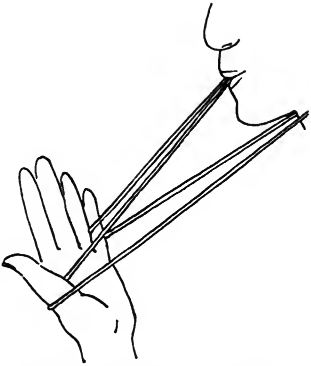


FIG. 15. A Step in "Dilly Bag."

left hand. Proceed with left index finger as in Pindiki and turn the left hand palm downward (Fig. 16).

4. *Right hand.* With the thumb and index finger pick up at their intersection the two strings which pass over the back of the left thumb.

5. *Left hand.* Withdraw hand from all strings. There are now two loops held by the right thumb and index finger. Hold left hand with fingers upward and insert the left thumb into the

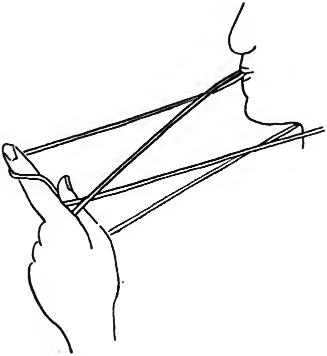


FIG. 16. A Step in "Dilly Bag."

near loop (left loop) and the little finger into the far loop (right loop).

6. Repeat Movements 3, 4 and 5.

7. Repeat again Movements 3, 4 and 5.

8. *Left hand.* Repeat Movement 3. Drop string from teeth (Fig. 17).

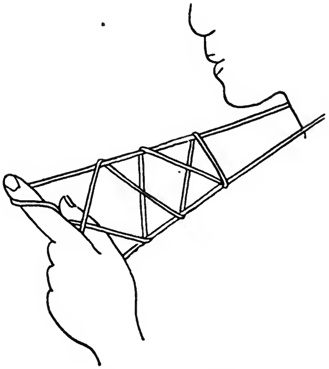


FIG. 17. Dilly Bag.

VII. BABY BEING BORN

WARDAMAN.

1. *Both hands.* First position.

2. *Right hand.* With index finger pick up from below the left palmar string and extend.

3. *Left hand.* Bend thumb over far thumb string and pick up from below on its back the near little finger string. Extend left hand. Proceed with left index finger as in Pindiki and turn the left hand palm downward (see left hand in Fig. 16).

4. *Right hand.* Bend middle, ring and little fingers downward over index finger strings and press against palm. With the thumb and index finger pick up at their intersection the two strings which pass over the back of the left thumb.

5. *Left hand.* Withdraw hand from all strings. There are now two loops held by right thumb and index finger. Hold left hand fingers upward and insert thumb into the near loop (left loop) and little finger into the far loop (right loop).

6. Repeat Movements 3, 4 and 5.

7. Repeat again Movements 3, 4 and 5.

8. *Left hand.* Repeat Movement 3. Move hand downward waist level.

9. *Right hand.* Move hand palm downward to a vertical osition over left hand. Hold all strings taut and drop index nger loop (Fig. 18). Shake the hand gently until the diagonal trings settle on the left hand (Fig. 19).

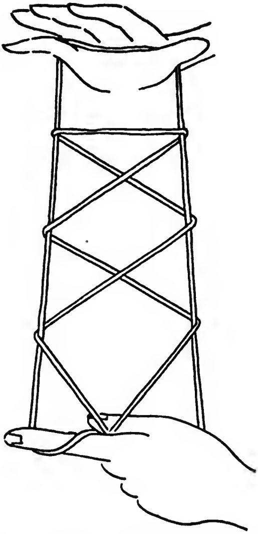


FIG. 18. A Step in "Baby Being Born."

VIII. BIG RIVER BAT

WARDAMAN.

1. *Both hands.* First position. Drop thumb loop.

2. *Left hand.* Move hand forward and to right around strings on right hand, thence back toward body under strings of right hand so that latter fall across left forearm.

3. *Right hand.* Move hand outward over left forearm then downward and inward through the loose loop hanging between

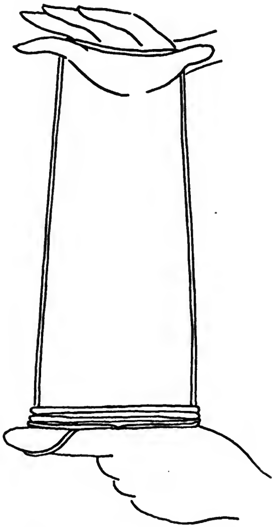


FIG. 19. Baby Being Born.

left little finger and left forearm to catch this loop on the back of the right forearm.

4. *Both hands.* Return to position by permitting arm strings to slide to wrist (Fig. 20).

5. *Both hands.* Bend thumb toward you over near wrist strings by raising elbow and turning palm downward. Pass

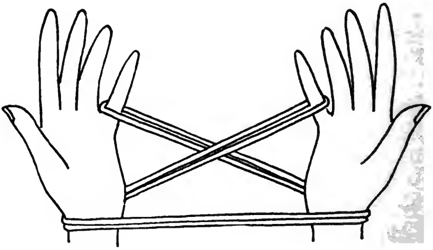


FIG. 20. A Step in "Big River Bat."

thumb outward and pick up from above on its back the far wrist strings. Draw these under near wrist strings. Return to position by lowering elbows.

6. *Right hand.* With thumb and index finger pick up strings on back of left wrist, lift over left hand and drop on palmar side.

7. *Left hand.* Repeat Movement 6 with left hand.

8. *Right hand.* Turn hand downward, palm outward, by raising elbow.

9. *Left hand.* Turn hand downward, palm inward, by raising elbow.

10. *Both hands.* Draw strings taut by extending hands (Fig. 21).

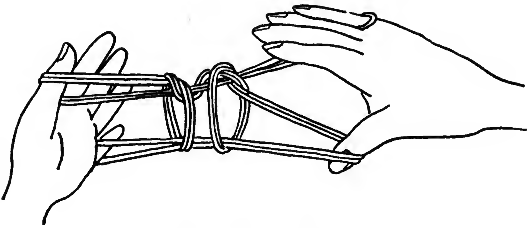


FIG. 21. Big River Bat.

Release

11. *Both hands.* Drop thumb strings, separate hands and the bat flies away.

IX. BLACKFELLOW

MELVILLE ISLAND, EAST OF DARWIN, NGAINMAN "Penis."

This figure is identical in manipulation with Yap "One Chief," D'Entrecasteaux "A Hiss," New Caledonia "Looper Caterpillar" and, except for very minor variations, Fiji "Ba-nuve." Comparisons should be made with Queensland "Giant Crane" and "Flying Fox," Arnhem Land "Little Bird" and figures from Palau and Nauru, the procedures of which are lacking.

1. *Both hands.* First position.

2. *Right hand.* With thumb and index finger pick up left far little finger string, bring toward you and wind once around

left little finger. Return to position. Insert right index finger from above into the little finger loop just formed. Turn the hand palm downward and outward so that picked up string becomes a loop on the right index finger. Extend.

3. *Left hand.* Pick up from below on back of index finger the right palmar string between the index finger strings. Extend.

4. *Right hand.* Repeat Movement 3 with right hand. Turn hand palm downward, fingers to left.

5. *Left hand.* Drop all strings. Insert thumb from left into upper (left) loop on right index finger and remove. Insert little finger into remaining loop on right index finger and remove.

6. *Both hands.* Hold fingers tightly together, turn hands fingers downward until central figure rests on knee. Drop thumb strings. Pass thumb downward into little finger loop, raise elbows slightly and, moving thumb outward, pick up on its ball the near little finger string (the farther string on your knee) and the nearer string of the far loop, the former far thumb string. Return to position.

(6A. *Both hands.* Variant manipulation: Insert index finger from below into thumb loop and withdraw thumb. Pass thumb under index finger strings and pick up from below the near little finger string. Keep thumb on far side of index finger. Bend index finger down over far index finger string and place tip against tip of thumb. Turn hand downward and permit index finger loop to slip on to thumb. Return to position.)

7. *Both hands.* Pindiki and draw out, working strings until figure is obtained (Fig. 22).

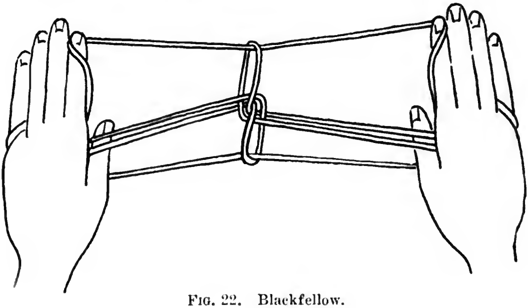


Fig. 22. Blackfellow.

X. LILY

WARDAMAN, NGAINMAN.

In the Diamantina District, South Australia, the pattern is known as "Yam."

1. *Both hands.* First position.

2. *Right hand.* Drop strings. Pick up from above with ball of index finger the left palmar string. Draw out and drop.

3. *Right hand.* Pick up from above with ball of index finger the left palmar string and draw out between strings of loop dropped.

4. *Left hand.* With thumb and index finger pick up right index finger loop and place over right wrist. Extend.

5. *Right hand.* Pass hand outward over far right wrist string, then leftward and downward, then inward under the far right wrist string, and upward until fingers touch left palm. Move fingers upward along palm under left palmar string. Insert little finger into left little finger loop and index finger into left thumb loop. Bend these fingers downward along sides of left hand to pick up from above the far left little finger string and the near left thumb string respectively. Draw out between the strings of the right wrist, elevate elbow to allow the right wrist loop to slip over the hand on to the strings drawn out. Extend (Fig. 23).

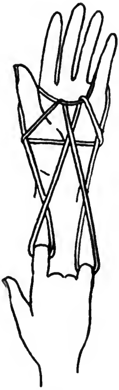


FIG. 23. A Step in "Lily."

6. *Right hand.* Hold strings taut and pass hand over left hand so that the string running from the near side of the left thumb passes between the left thumb and index finger, the string running from the far side of the left little finger passes between the left little finger and ring finger, and the two crossed strings pass respectively between the left index and middle fingers and between the left middle and ring fingers. Drop strings from right hand over back of left hand (Fig. 24).

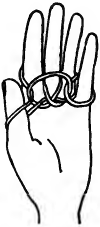


FIG. 24. A Step in "Lily."

7. *Right hand.* With thumb and index finger pick up loose palmar string of left hand and draw out (Fig. 25). Drop this string. With thumb and index finger pick up the intersection of all strings.

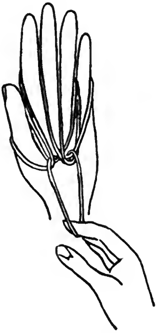


FIG. 25. A Step in "Lily."

8. *Left hand.* Withdraw hand from all strings (Fig. 26).

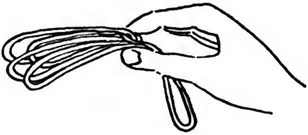


FIG. 26. Lily.

XI. HEADDRESS

NGALUMA.

This figure is similar to Lily (X) but there are some interesting variations.

1. *Both hands.* First position.

2. *Right hand.* Drop strings. With thumb and index finger pick up left thumb loop, lift off thumb and place on left index finger. With ball of index finger pick up from above the left palmar string. Draw out and drop.

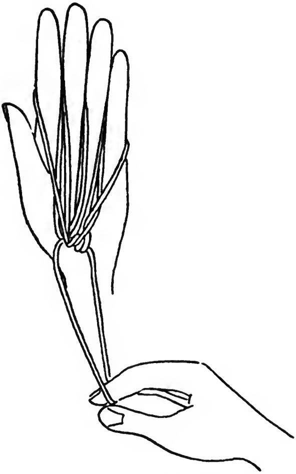


FIG. 27. Headdress.

3. *Right hand.* With ball of index finger pick up from above the left palmar string and draw out between strings of loop just dropped. Insert from above the right thumb, middle, ring and little fingers into loop held by ball of right index finger and move hand to the left until tips of fingers touch left palm. Insert little finger under left far little finger string and index finger under near left index finger string. Bend these fingers downward along sides of left hand to pick up the strings indicated and draw out through long loop.

4. *Right hand.* Hold strings taut and pass hand over left hand so that the near left index finger string passes between the left index and middle fingers, the far left little finger string passes between the left little and ring fingers and the other two strings pass between the left middle and ring fingers.

5. *Right hand.* Drop strings of right hand over back of left hand. With thumb and index finger pick up the loose palmar string of left hand and draw out (Fig. 27).

XII. HUT

KELLERBERRIN, SOUTHERN CROSS, NGALUMA, NANGUMARDA, TAGUDI, INJIBANDI, PANJIMA, TALAINJI, BAIONG, INGARDA "Penis," KIDJA, DJARU "Yam."

1. *Both hands.* First position.

2. *Right hand.* Pick up from below with index finger the left palmar string and draw out dropping right thumb and little finger strings. Drop index finger loop and allow it to hang downward. Pick up with index finger from below the left palmar string between the strings of the loop just dropped. Draw out and drop.

3. *Left hand.* Hold hand palm downward with fingers toward the right.

4. *Right hand.* Pass right hand into the long loop from the right. Turn fingers upward and with the index finger pick up from below the near left thumb string. In a similar manner pick up from below the far left little finger string with the right little finger. Draw these two strings out through the long pendant loop.

5. *Left hand.* Turn hand so that fingers point upward. Hold strings taut and bend the index, middle and ring fingers downward so that index finger passes into loop held by right

ex finger; the ring finger into loop held by right little finger; middle finger between these two loops.

6. *Right hand.* Pass right hand over left hand and drop all ht hand strings over the back of the left hand.

7. *Left hand.* Open hand and point fingers upward.

8. *Right hand.* With thumb and index finger pick up loose ing on left hand which passes over the other strings and draw t. Move hand over left hand and turn the palm of the latter ward (Fig. 28).

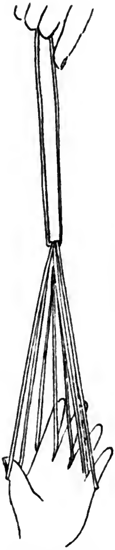


FIG. 28. Hut.

XIII. EMU FOOT

uairading, Gnowangerup.

This figure is similar to Central Australian "Rabbit-Bandioot" and "Two Dogs," the procedures of which are not given.

1. *Both hands.* First position.

2. *Right hand.* Pick up from below with the index finger the eft palmar string, draw out slightly and rotate the finger one

complete turn to the right, thus twisting the right index strings. Extend.

3. *Left hand.* Pick up from below on the back of the index finger the right palmar string and draw out between the strings of the right index loop.

4. *Right hand.* Drop thumb and little finger strings and extend (Fig. 29).

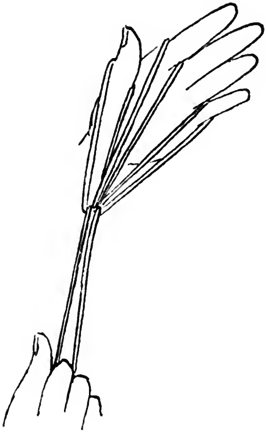


FIG. 29. Emu Foot.

XIV. EMU

SOUTHERN CROSS, LAVERTON.

This figure is one of the most widely distributed in the world but the procedure differs greatly in most regions.

Two persons are required for this figure.

1. *Both hands.* First Position. Drop little finger strings.

2. *Right hand.* Give thumb loop to second person to hold, strings taut across body. Place hand over strings about a foot from loop held by second person and hook little finger downward around near string. Bend index finger downward and hook far string, then draw it back over string held by little finger by turning elbow outward. Pass index finger downward and straighten

thus catching on back the little finger string. Rotate hand to right and loosen strings slightly by moving hands toward each other, until it is possible to insert index finger from above into loop now found on little finger. Continue rotating movement of index finger and catch on it as you return the far string of the loop held by second person. Draw this string toward you as second person drops loop and extend.

3. *Left hand.* Proceed as in Movement 2, but reverse direction of rotation (Fig. 30).

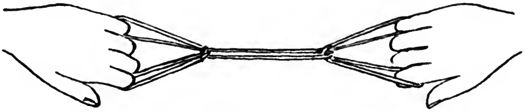


FIG. 30. Emu.

Release normal

STRING FIGURES FORMED FROM OPENING A

XV. TWO KANGAROOS

WARDAMAN, NGAINMAN.

This figure is practically identical in procedure with Yap "No Name," Gilbert Islands "Kani Mumum" and, except for the intermediate steps to form "Te Tuma O Te Papa," with Society Islands "Vahia." The finished figure is comparable with one of the Marind-anim (New Guinea) figures.

1. *Both hands.* Opening A. Pass little, ring and middle fingers and thumb from below into index finger loop which thus becomes a wrist loop.

2. *Both hands.* Insert index finger from above into little finger loop, bend it toward you and insert from above into thumb loop. Pick up with back of index finger the far thumb string and draw it up through the little finger loop. Extend.

3. *Both hands.* Drop thumb loop. Insert thumb from below into index finger loop and withdraw index finger.

4. *Right hand.* Hold strings secure, reach over back of left hand and with thumb and index finger pick up left wrist string, lift off left hand and place on left thumb and index finger as in First position.

5. *Left hand.* Repeat Movement 4 with left hand.

6. *Both hands.* Repeat Movement 2.

7. *Both hands.* Hold strings slack and drop little finger loops. Extend. Reach little finger forward and pick up from below with its back the near index finger string. Return to position and drop index finger strings.

8. *Both hands.* There are now two loops on the thumb and one on the little finger. On each palmar string is a loop one string of which runs directly across figure to other palmar string. Bend the index finger down over palmar string and with its ball pick up from the far side this horizontal string. Return to position by rotating the hands toward you and lowering slightly the elbows, thus placing a loop on the back of first joint of the index finger.

9. *Both hands.* Drop thumb loops. Turn palms downward and draw out working strings slowly until figure is obtained (Fig. 31).

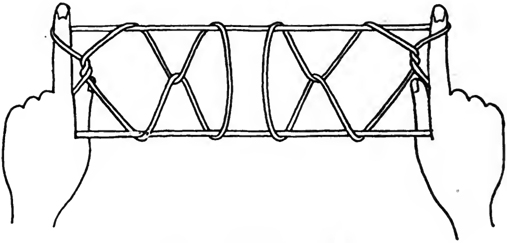


FIG. 31. Two Kangaroos.

Release normal

XVI. TWO TURTLES ON A LOG

MARAURA.

Procedure is identical with South Australia "Two Swans."

1. *Both hands.* Opening A.

2. *Both hands.* Bend thumb over index finger strings and pick up from below with its back the near little finger string. Extend.

3. *Both hands.* Bend index finger down over palmar string and insert between strings of its own loop. Press tip of finger

against palm of hand and release the little finger loop and at the same time draw strings taut and turn hands palms outward. Straighten index finger and return to position.

4. *Both hands.* Insert middle, ring and little fingers from below into thumb loops. Straighten the middle finger slightly to lift up on its back the far lower thumb string. Bend the ring and little fingers down over the near upper thumb string and pick up from below with the back of the little finger the near lower thumb string. Return to position by straightening the little finger to draw out the string picked up and by withdrawing the middle and ring fingers.

5. *Both hands.* Drop thumb strings and turn hands fingers outward, palms facing. Pass thumb from below into little finger loop. Bend index finger downward, place tip against tip of thumb and allow the index finger loop to slip on to thumb. Retain the near little finger string also on thumb and return to position. Pindiki and extend (Fig. 32).

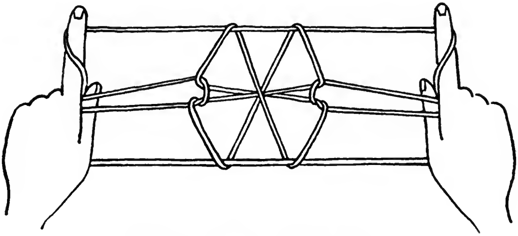


FIG. 32. Two Turtles on a Log.

Release normal

XVII. KANGAROO TRAIL

WARDAMAN.

1. *Both hands.* Opening A.

2. *Both hands.* Bend thumb over index finger strings and pick up from below the near little finger string.

3. *Both hands.* Bend middle finger over palmar and index finger strings and pick up from below the far thumb string. Extend.

4. *Both hands.* Drop little finger loop. Extend. Bend little finger forward, insert from below into middle finger loop and withdraw middle finger. Extend.

5. *Both hands.* Repeat Movements 3 and 4.

6. *Both hands.* Bend thumb down over index finger strings, draw hands slightly apart, thus releasing upper thumb loop. Extend.

7. *Both hands.* Turn hands palms inward, fingers upward. Bend all fingers over index finger strings and insert from below into thumb loop. Extend fingers to change thumb loop to a wrist loop.

8. *Right hand.* With tips of thumb and index finger pick up left near little finger string and the left far index finger string, close to left palm.

9. *Left hand.* Remove from all strings, permitting wrist string to hang downward. With palm downward, fingers outward, move hand toward right hand and insert thumb into hanging loop. Bring thumb toward you by moving elbow outward to pick up loop on back of thumb. Lower elbow and turn hand fingers upward. Insert index finger upward into near loop held between right thumb and index finger, and the left little finger similarly into far loop. Withdraw these loops from right hand.

10. *Left hand.* Repeat Movement 8 with left hand.

11. *Right hand.* Repeat Movement 9 with right hand. Extend.

12. *Both hands.* Reach thumb over index finger strings and

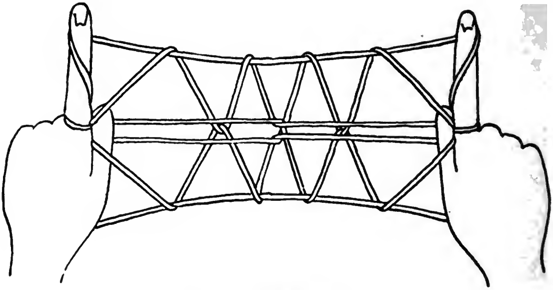


FIG. 33. Kangaroo Trail.

Release normal

pick up from below the near little finger string. Pindiki and extend, slowly working strings until figure is obtained (Fig. 33).

XVIII. Dead Kangaroos

A continuation of Kangaroo Trail (XVII)

Wardaman.

The procedure differs from D'Entrecasteaux "The Seine" and Gilbert Islands "Te Kananei."

1. *Both hands.* Construct "Kangaroo Trail" (XVII).

2. *Both hands.* Drop index finger strings, draw hands slowly apart, strings loosely held. Turn palms facing, fingers downward and place right hand figure of "kangaroo" across right knee.

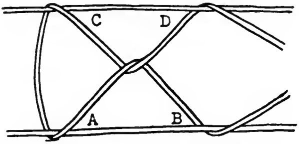


Fig. 34. A Step in "Dead Kangaroos."

3. *Left hand.* Withdraw hand from all strings and let them fall across left knee.

4. *Right hand.* On the right knee are two parallel horizontal strings which hold the figure of the "kangaroo" (Fig. 34). The center of the latter consists of an intersection of four strings, two of which (*A* and *C*) pass over the near and far horizontal strings respectively to meet at the intersection whence their continuations (*B* and *D*) go back and under the respective hori-

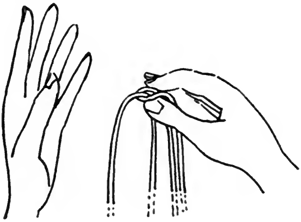


Fig. 35. A Step in "Dead Kangaroos."

zontal strings. Remove right hand from all strings and pick up the intersection with the tips of thumb and index finger, the thumb to the left of the intersection, the index finger to the right. Lift intersection a few inches and turn hand so that fingers point to the left (Fig. 35).

5. *Left hand.* Bring hand toward the right and pick up on the back of the little finger the string on the far side of the intersection (*D*) and on the back of the thumb the string on the near side (*C*). Grasp the intersection held in right hand between the tips of the left thumb and index finger.

6. *Right hand.* Release intersection from right hand and similarly to Movement 5 pick up respectively on thumb and little finger the near (*A*) and far (*B*) strings of the intersection held in left hand.

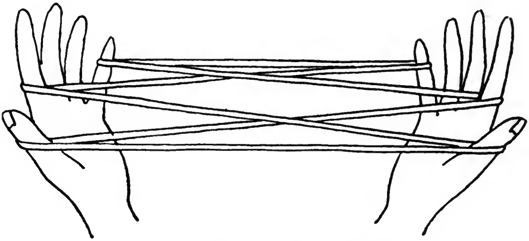


FIG. 36. A Step in "Dead Kangaroos."

7. *Left hand.* Drop intersection held by thumb and index finger and hold hands closely together.

8. *Right hand.* Bend index finger over the right near little finger string and by passing the finger downward then inward pick up from the far side the left near little finger string.

9. *Left hand.* Bend index finger downward over the left far thumb string and pick up from below on its back the far right thumb string.

10. *Both hands.* Extend. The strings are now as in Opening A with the exception that the strings comprising the index finger loops are crossed, whereas in Opening A one set has been drawn through the other (Fig. 36).

11. Proceed as in "Kangaroo Trail," Movements 2 to 13.

12. *Both hands.* Drop both index finger loops, Pindiki and extend slowly working the strings until figure is obtained (Fig. 37).

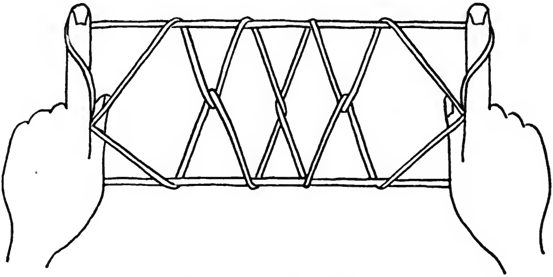


FIG. 37. Dead Kangaroos.

Release normal

XIX. BIG MOB OF BLACKFELLOWS

WARDAMAN.

Compare finished figure with Queensland "Two Tomahawks" and "Rain."

1. *Both hands.* Opening A.

2. *Both hands.* Insert little, ring and middle fingers and thumb into index finger loop from below and let it slide over back of wrist. Pick up from below with thumb the near little finger string and return to position.

3. *Left hand.* Insert index finger from below into thumb loop and pick up the far thumb string on back of first joint. Press thumb against base of index finger and turn hand palm to right.

4. *Right hand.* Hold strings secure and with tips of thumb and index finger grasp the two strings which pass over back of left thumb.

5. *Left hand.* Withdraw little finger and thumb from their loops but retain loop on wrist. Move hand palm downward, fingers to right, toward the right hand and pick up on little finger the far loop held by right thumb and index finger. Move the left hand slightly away from you. Bend the thumb downward and then, by moving the hand toward you, pick up with the thumb the remaining (near) loop.

6. *Right hand.* Repeat Movement 3 with right hand.

7. *Left hand.* Repeat Movement 4 with left hand.

8. *Right hand.* Repeat Movement 5 with right hand.

9. *Both hands.* Pick up from below with back of thumb the near little finger string and return to position.

10. Repeat Movements 3 to 9.

11. Repeat again Movements 3 to 9.

12. *Right hand.* Hold all strings secure, reach over left hand, grasp left wrist string between tips of thumb and index finger, lift over left hand and let fall between hands.

13. *Left hand.* Repeat Movement 12 with left hand.

14. *Both hands.* Pindiki and work strings slowly until figure is obtained (Fig. 38).

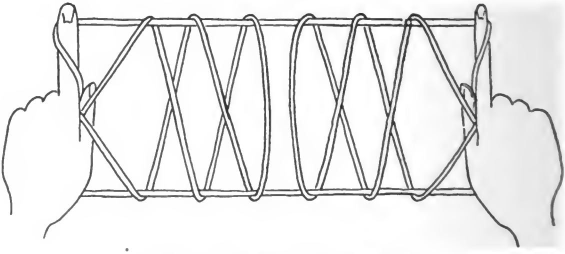


Fig. 38. Big Mob of Blackfellows.

Release normal

XX. Armband

Wardaman.

Except for minor variation the procedure for this figure is identical with that for Yap "Three Stars."

1. *Both hands.* Opening A. Pass little, ring and middle fingers and thumb into index finger loop and permit it to slide to wrist.

2. *Left hand.* Insert index finger from above into thumb loop, then reach away from you and insert from above into little finger loop. Pick up the near little finger string with the ball of the index finger and draw up through the thumb strings by rotating the index finger upward and turning the palms upward.

3. *Right hand.* With the tips of the thumb and index finger pick up near the base of the left index finger the intersection formed by the left near index finger string and the left far thumb string.

4. *Left hand.* Withdraw thumb, index finger and little finger from all strings but retain the wrist loop. Pass the hand over the right hand and moving it back toward the left pick up from the right with the little finger the near loop held by the right thumb and index finger (previously the left thumb loop). Hold this loop secure by bending little finger against palm. Turn hand palm downward by elevating the elbow. Again place hand over right hand and move it to the left to pick up from the right on the ball of the index finger the far loop held by the right hand (previously the left index finger loop). Rotate the index finger downward, then toward you, and extend the index finger. Straighten the little finger. Pass the thumb from below into index finger loop and withdraw index finger.

5. *Right hand.* Repeat Movement 2 with right hand.

6. *Left hand.* Repeat Movement 3 with left hand.

7. *Right hand.* Repeat Movement 4 with right hand.

8. Repeat Movements 2 to 7.

9. Repeat again Movements 2 to 7.

10. *Left hand.* Repeat Movement 2. Turn hand palm upwards and extend.

11. *Right hand.* With tips of thumb and index finger pick up the left wrist string near the base of the left thumb, lift over thumb and lay across the intersection of the left far thumb string and the left near index finger string. With the tips of the right thumb and index finger pick up the intersection of these three strings.

12. *Left hand.* Withdraw hand from all strings. Pick up with the tips of the thumb and index finger the middle string held by the right thumb and index finger (the loop moved in Movement 11). Bend this string down tightly between the other two loops and place between right thumb and index finger from the left. This movement leaves just the two original loops appearing above the tips of the right thumb and index finger.

13. *Left hand.* Hold hand palm downward, fingers outward, and pass over and to the right of the right hand. Move the hand back toward the left and by sliding the little finger along the

right thumb pick up the near loop held by tips of right thumb and index finger. Raise the elbow and bend the left thumb downward so that it can be inserted from below through the loop just picked up on little finger. Pass thumb onward into the far loop held by tip of right thumb and index finger and return left hand to position by taking these strings from the right hand.

14. *Right hand.* Repeat Movement 2 with right hand. Turn hand palm upward and extend.

15. *Left hand.* Repeat Movement 11 with left hand.

16. *Right hand.* Repeat Movement 12 with right hand.

17. *Right hand.* Repeat Movement 13 with right hand.

18. *Both hands.* Return to position, Pindiki and extend (Fig. 39).

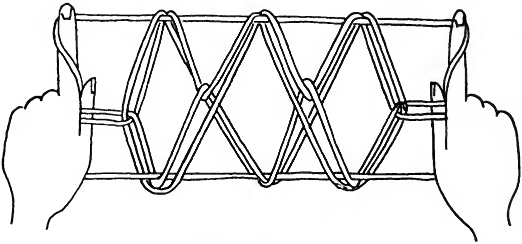


FIG. 39. Armband.

Release normal

XXI. FISH SPEAR

WARDAMAN.

The procedure in this figure is similar to that in several other Australian figures, such as North Australian "Lily," Queensland "A Lily Root," Northwestern "Headdress," Central Australian "Cooked Rabbit on a Stick," etc.

1. *Both hands.* Opening A. Take up left palmar string first.

2. *Right hand.* Drop index finger loop and extend.

3. *Left hand.* Bend index finger down over palmar string and into its own loop. Press tip against palm to secure palmar string, raise elbow and drop all other strings on hand. Extend and drop index finger loop.

4. *Right hand.* Turn hand palm downward, fingers to left.

5. *Left hand.* Pass hand from left into long loop hanging from right hand. Move fingers along base of right hand and pick up from below on index finger and little finger respectively the near right thumb string and the far right little finger string. Draw out through long, pendant loop and extend.

6. *Right hand.* Turn hand fingers upward. Hold strings taut and bend index, middle and ring fingers downward in such a way that index finger passes into left index finger loop, ring finger into left little finger loop and middle finger between the strings of the two loops respectively.

7. *Left hand.* Pass hand over right hand and move to right until strings are taut. Drop left hand strings over back of right hand.

8. *Right hand.* Open hand, fingers upward.

9. *Left hand.* With index finger pick up from above the loose palmar string on right hand and draw out (Fig. 40).

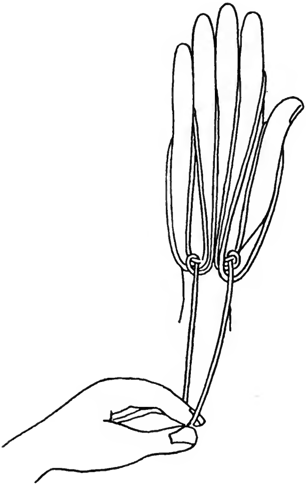


Fig. 40. Fish Spear.

XXII. VULVA

An action figure

WARDAMAN.

1. *Right hand.* Place loop under foot and draw out into a long loop.

2. *Both hands.* Insert hands into loop as in First Position and proceed as in Opening A.

3. *Both hands.* Pass thumb over index finger strings and pick up from below the near little finger string, but do not return to position.

4. *Teeth.* Pick up middle of lower near thumb string between the teeth and draw out.

5. *Both hands.* Draw the upper thumb string under the string held by teeth to Navaho the two loops. Drop string from teeth. Action is indicated by drawing hands apart, then alternately drawing them upward away from foot (Fig. 41).

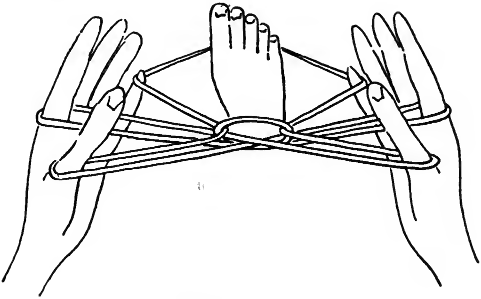


FIG. 41. Vulva.

XXIII. SPEAR

ARUNTA.

1. *Both hands.* Opening A. Take up left palmar string first.

2. *Right hand.* Drop index finger loop and extend. Drop thumb and little finger loops and let them hang below hand. With thumb pick up from the near side the near and far left index finger strings below the left palmar string. Draw out.

With the back of the little finger pick up from below the far thumb strings as in First position.

3. *Left hand.* With index finger pick up from below the right palmar strings as in Opening A. Hold these strings secure by pressure from the thumb and middle finger.

4. *Right hand.* Drop all strings. Pass hand over left hand and pick up with tips of thumb and index finger the lower left index finger loop. Draw upward and rightward over the upper left index finger loops and extend (Fig. 42).

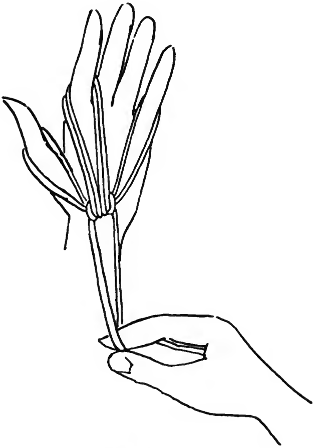


FIG. 42. Spear.

XXIV. BED

SOUTHERN CROSS, GNOWANGERUP, MOORA.

Similar patterns from different procedures are found in "Goanna" (XXXIX) and "Bunk" (XL).

1. *Both hands.* Opening A.
2. *Both hands.* Drop little finger strings and extend.
3. *Both hands.* Pass little finger over index finger strings and pick up from below the far thumb string. Return to position and turn fingers upward.

4. *Right hand.* Pick up left index finger loop, lift off finger and place over all fingers of left hand to become a wrist loop.

5. *Left hand.* Repeat Movement 4 with left hand.

6. *Right hand.* With thumb and index finger grasp the far thumb and near little finger strings of left hand.

7. *Left hand.* Withdraw hand from all strings, then move hand to right and insert little finger and index finger into the far and near loops respectively held by right thumb and index finger.

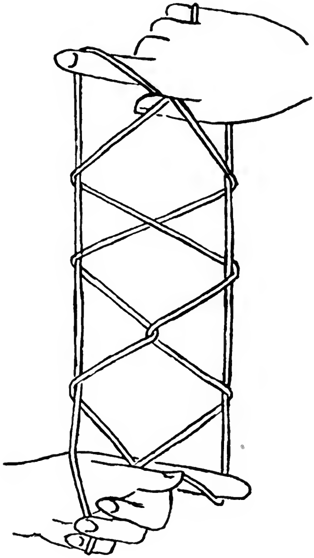


Fig. 43. Bed.

8. *Left hand.* Repeat Movement 6.

9. *Right hand.* Repeat Movement 7. Extend.

10. *Both hands.* Turn hand fingers outward. Bend thumb downward and pick up from below the loop which runs around all parallel strings. Drop index finger loop and extend.

11. *Both hands.* Pick up from below with thumb the near little finger string. Pindiki and extend slowly as you bring right

hand upward in front of face, then downward and outward, at the same time bringing the left hand inward and upward (Fig. 43).

XXV. Bunk

Djaru, Kidja, Nangumarda.

Several Central Australian patterns are identical.

1. *Both hands.* Opening A.

2. *Both hands.* Drop thumb loop. Pass thumb over near index finger string, pick up from below near little finger string and return to position.

3. *Both hands.* Bend down middle finger into index finger loop, pick up from below the near index finger string (on central side of far thumb string) and return to position. Drop little finger string so that there is a long pendant loop, but do not extend.

4. *Both hands.* Move hands slightly together. Withdraw thumb from its loop, reach thumb over it and pick up on back of thumb the long pendant loop (formerly the far little finger string). Draw out and extend (Fig. 44).

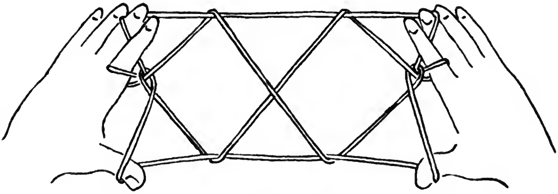


Fig. 44. Bunk.

Release normal

XXVI. Eaglehawk

Wardaman.

Two persons and a long string are required for this figure.

1. *First Person. Both hands.* Opening A.

2. *Second Person. Both hands.* Face First Person and insert hands from below into little finger loop. Draw his far little

finger string slightly toward you. Pass little finger away from you over his near little finger string and under his index finger strings. Turn palms upward and move hands together until little fingers touch. On each little finger there will be two strings (the index finger strings of the First Person).

3. *First Person. Both hands.* Drop all strings except thumb loop. Pass all fingers from below into thumb loop to change it to a wrist loop. Reach over palms of Second Person and pick up with tips of thumb and index finger his lower little finger string (the one farther from you), lift up and place over his thumb. With tips of thumb and index finger pick up the remaining little finger string and draw slightly toward you.

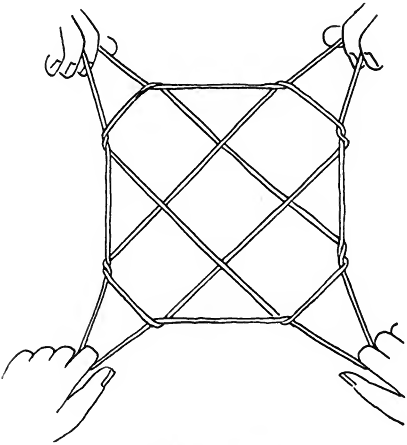


Fig. 45. Eaglehawk.

4. *Second Person. Both hands.* Drop all strings except thumb and wrist loops. Let loop on thumb slip to its tip and hold fast with tip of index finger.

5. *Both Persons. Both hands.* Pull string held by thumb and index finger toward you, raise elbows to permit wrist loop to slip over hand and on to strings, and extend hands slightly. Shift loop held by thumb and index to a wrist loop by passing into it from below all fingers and thumb.

6. *First Person. Both hands.* Bend hands downward into large triangle formed by your near wrist strings and their continuation. Hold the latter side of the triangle on the backs of the fingers and grasp between the tips of thumb and index finger the corresponding string of the Second Person.

7. *Second Person. Both hands.* Reach forward with tips of thumb and index finger and pick up the string resting on the backs of the fingers of the First Person.

8. *Both Persons. Both hands.* Raise elbows to change string held by thumb and index finger to the ball of the index finger, draw hands slightly apart and toward you and permit wrist loop to slip easily over hand on to loop drawn out (Fig. 45).

Release normal

XXVII. Big Pig

Wardaman.

Two persons and an extra long string are required.

1. *Both Persons.* Stand face to face and pick up string in First position so that a rectangle is formed. Opening A (Fig. 46).

2. *First Person. Right hand.* With tips of thumb and index finger pick up left index finger loop and place as a wrist loop over right hand of Second Person (Fig. 47).

3. *First Person. Left hand.* Repeat Movement 2 with left hand.

4. *Second Person. Right hand.* Repeat Movement 2.

5. *Second Person. Left hand.* Repeat Movement 2 with left hand (Fig. 48).

6. *Both Persons. Right hand.* Reach over left palm and lift up between tips of thumb and index finger the left far thumb string and the left near little finger string.

7. *Both Persons. Left hand.* Withdraw hand from all strings, then pass hand to right and pick up on the little finger and thumb respectively the far and near loops held by right thumb and index finger.

8. *Both Persons. Left hand.* Repeat Movement 6 with left hand.

9. *Both Persons. Right hand.* Repeat Movement 7 with right hand.

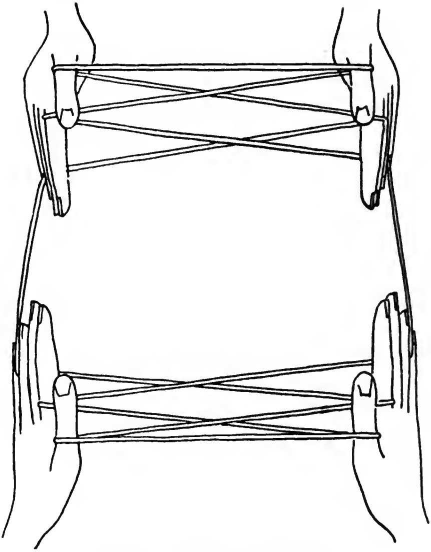


FIG. 46. A Step in "Big Pig."

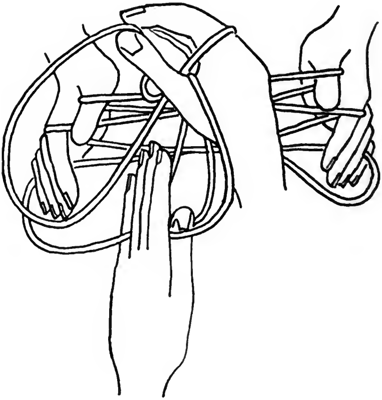


FIG. 47. A Step in "Big Pig."

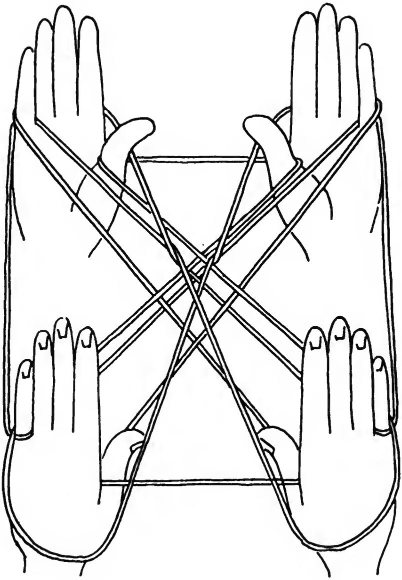


FIG. 48. A Step in "Big Pig."

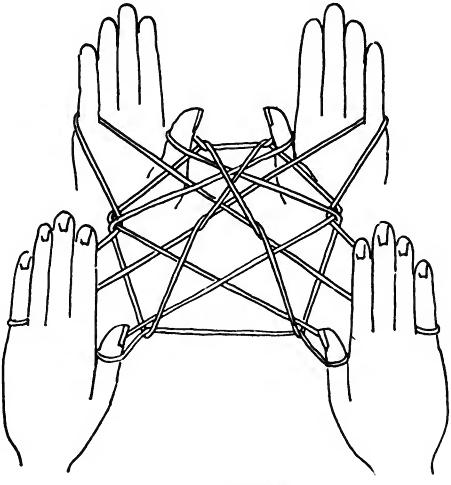


FIG. 49. Big Pig.

10. *First Person. Both hands.* Hold hands fairly close together, fingers joined and upward.

11. *Second Person. Both hands.* Draw hands apart, fingers upward and well extended (Fig. 49).

Release normal

XXVIII. Dilly Bag

Wardaman.

"Tortoise" in both Arnhem Land and Queensland is similar in pattern. Two persons are required for this figure.

1. *Second Person.* Opening A. Face First Person. The strings are named from point of view of Second Person although First Person performs all subsequent manipulations.

2. *First Person. Both hands.* Pick up near thumb string and place in mouth of Second Person. Reach over mouth string and grasp between tips of thumb and index finger either of the diagonal strings running from far side of thumb to near side of index finger of other hand. Pull this string up over mouth string, then downward to form a small loop. Pass hand from above into this loop and grasp the other diagonal string. Draw out through loop, then downward to form a new loop. Proceed similarly by drawing through each new loop, first one, then the other diagonal string which runs from far side of index finger to near side of little finger of opposite hand. Finally pick up the far little finger string, draw out into a loop and place over

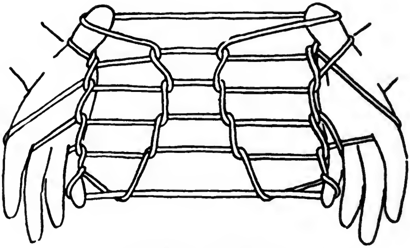


Fig. 50. Dilly Bag.

little finger of Second Person. Pick up loop held in mouth of Second Person, extend and place over his thumbs (Fig. 50).

Release normal

XXIX. TURTLE

WAGOMAN, KIDJA, DJARU.

Arnhem Land "Tortoise" is similar in pattern. Procedure is not given. Two persons and a long string are required.

1. *First Person. Both hands.* Opening A. Turn fingers outward. The strings are named from point of view of First Person although Second Person performs all subsequent manipulations.

2. *Second Person.* With right thumb and index finger pull out the far little finger string between left thumb and index finger to form a loop. Hold loop secure with left hand, insert right thumb and index finger through it from below and grasp the diagonal string which runs from the right little finger to the left index finger. Draw out through loop and form a new loop. Proceed in the same manner with the other diagonal string and in their turn with the strings running between each thumb and the opposite index finger. Place the last loop in teeth of First Person.

3. *Second Person.* Place the loops held by left thumb and index finger of own hand on middle finger. Proceed as in Movement 2 by drawing out through the left thumb and index finger the bottom string to the right of the strings held in own left hand, then the other strings in the same order as before. Place last loop over the left thumb of First Person.

4. *Second Person.* Place the loops held by own left thumb and index finger on index finger. Proceed as in Movement 2, this time taking the strings in the same order to the left of the strings held in own left hand and drawing out the first loop between the left ring and little fingers. Place last loop over right thumb of First Person.

5. *Second Person.* With right thumb and index finger remove loop from teeth and place over thumbs of First Person. Place loops held by left ring and little fingers on left ring finger (Fig. 51).

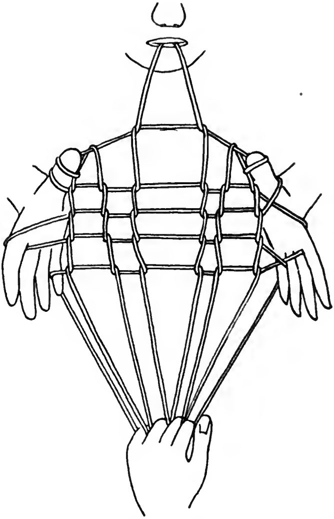


FIG. 51. Turtle.

XXX. TURTLE

DJARU, KIDJA.

Two persons are required for this figure.

1. *First Person. Both hands.* Opening A. Face Second Person with palms inward, fingers upward.

2. *Second Person. Both hands.* With palms facing pass hands from below into the little finger loops of First Person. Bend little finger downward and pass under index finger strings. Move hands together until sides of little fingers touch. Turn palms upward so that there will be two palmar strings across each little finger.

3. *First Person. Both hands.* Drop all strings except thumb loops. Pass all fingers from below into thumb loop and permit it to slip on to wrist. With tips of thumb and index finger pick up the upper palmar little finger string of Second Person (the string nearer to you) and draw out slightly.

4. *Second Person. Both hands.* Hook little finger over remaining palmar little finger string and draw out slightly. Place tip of index finger against tip of little finger, turn hand palm downward and hook string on index finger.

5. *Both Persons. Both hands.* Draw out to permit wrist loops to slip over back of hands (Fig. 52).

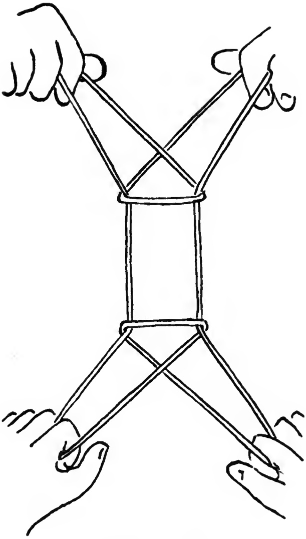


Fig. 52. Turtle.

Release normal

STRING FIGURES FORMED FROM OPENING A, MOVEMENT Z

Movement Z

A manipulation found in several figures from North Australia is herein called, for want of a better name, Movement Z. In so far as we know, it is an elaboration of Opening A but since it conceivably could be employed in other steps of construction some separate term seems warranted. An example of how the same result can be attained by a less dextrous manipulation is found in Movements 2–4 in "Gate" (XXXVII).

1. *Both hands.* Opening A.

2. *Right hand.* Pass thumb over all strings and press them downward. Reach downward over bent thumb with index finger and pick up from the far side with its ball the far little finger string. Lift this string about an inch over the back of the thumb (Fig. 53). Turn right palm toward you and at the same time extend thumb and index finger so that the picked up string will form a loop on their backs. Turn hand palm outward.

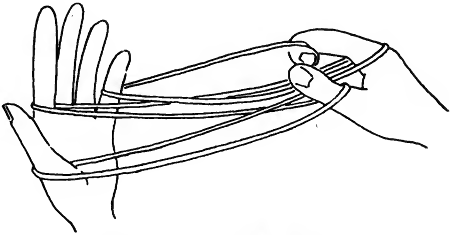


FIG. 53. A Step in Movement Z.

3. *Left hand.* Insert index finger from below into upper right thumb loop between right thumb and index finger and take loop off right thumb. Extend and return to position (Fig. 54).

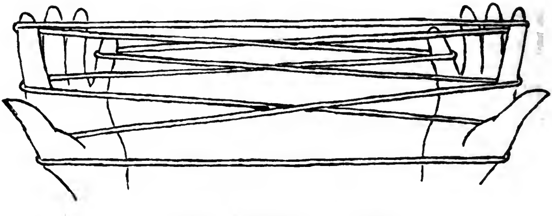


FIG. 54. Movement Z Completed.

XXXI. Two Blackfellows

Wardaman.

1. *Both hands.* Opening A. Movement Z.

2. *Both hands.* Insert little, ring and middle fingers and thumb from below into index finger loops and allow them to slide over back of hand.

3. *Both hands.* Insert index finger from above into little finger loop, bend it back toward you and insert from above into thumb loop. Pick up with back of index finger the far thumb string and draw it up through the little finger loop. Extend.

4. *Right hand.* Hold strings secure, reach over back of left hand and with thumb and index finger pick up left wrist strings, lift over left hand and place on left thumb and index finger as in First position.

5. *Left hand.* Repeat Movement 4 with left hand.

6. *Both hands.* Repeat Movement 3.

7. *Both hands.* Hold strings slack and drop little finger strings. Extend. Reach little finger forward and pick up from below with its back the near index finger string. Return to position and drop index finger strings.

8. *Both hands.* There are now three loops on each thumb, one on each little finger. From each palmar string a diagonal string crosses to the far side of the thumb of the opposite hand. Bend index finger down over the palmar string and on far side of the diagonal string. Lift this string slightly with the near side of the index finger close to the point where it intercepts the palmar string. Move the index finger toward you over the string which runs from the far side of the thumb to intercept the middle of the far little finger string and insert downward on the far side of the far thumb string which runs to the palmar string of the opposite hand (now supported by the index finger of the opposite hand). Pick up this string on the ball of the index finger, in addition to the one already caught there and return to position by rotating the hands toward you and lowering the elbows. This movement places on the first joint of each index finger two strings which now run directly from one index finger to the other.

9. *Both hands.* Drop thumb loops, turn palms outward and

draw out, working the strings slowly until figure is obtained (Fig. 55).

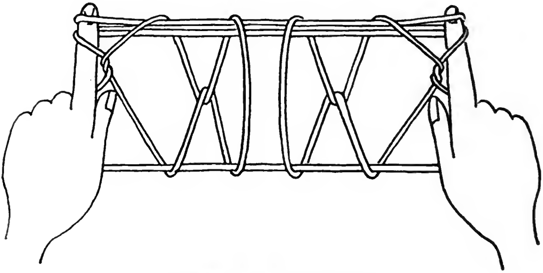


FIG. 55. Two Blackfellows.

XXXII. BAMBOO

WARDAMAN, NGAINMAN.

A long string is required for this figure. The procedure varies only in minor details from Yap "House of the Blos Bird" and Society Islands "Te Fare No Oro."

1. *Both hands.* Opening A. Movement Z.

2. *Both hands.* Movements 2 to 8 of "Two Blackfellows."

3. *Both hands.* Press middle fingers against far index finger strings to hold them secure. Drop thumb strings but do not extend. Tilt hands slightly outward so that strings just dropped

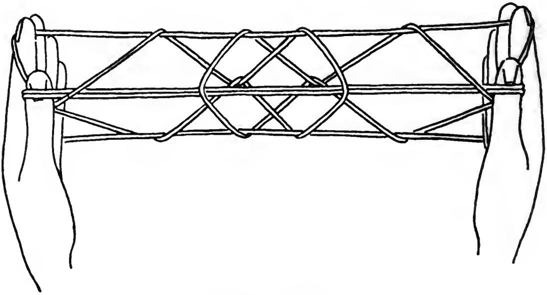


FIG. 56. Bamboo.

swing between palms. Reach thumb upward and place tip against tip of index finger. Return hands to position and tilt slightly toward you so that index finger loops slide on to thumb. Release pressure of middle finger and return to position.

4. *Both hands.* Pindiki and draw out, slowly working strings until figure is obtained (Fig. 56).

Release normal

XXXIII. Rain

Wardaman, Ngainman.

1. *Both hands.* Opening A. Movement Z.

2. *Both hands.* Drop thumb loop. Reach thumb over lower near index finger string and pick up from below on its back the two far index finger strings and draw out slightly between the upper and lower near index finger strings, palms toward you.

3. *Both hands.* Bend little finger over the upper near index finger string, insert it downward and pick up from the far side the lower near index finger string. Draw out slightly over the upper near index finger string and slide loop to base of little finger.

4. *Both hands.* Hold strings slack, palms toward you. Drop thumb loops. Reach thumb over strings just dropped and pick up from below on its back the upper near index finger string at the point where it runs across palm of little finger. Bend little, ring, middle and index fingers down over palmar string and between little and index finger strings and press tips against palm. Do not permit little finger strings to slip past the second joint. Raise elbows slightly and extend carefully (Fig. 57). (A well spaced final figure is difficult to attain.)

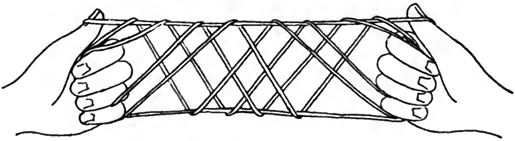


Fig. 57. Rain.

Release normal

XXXIV. Devil-devil's Anus

Ngainman.

This figure is an intermediate step in the construction of "Crocodile" (XXXV), one of the most widely distributed figures in the Pacific Islands and Australia. It is seldom, however, that this step is considered a separate figure and given a specific name. It is therefore important to note this recognition in North Australia, Queensland "Mouth of a Fish," and Fiji "Vasua."

1. *Both hands.* Opening A. Movement Z.

2. *Both hands.* Drop thumb loop. Extend. Pass thumb under index finger strings, pick up from below the near little finger string on its back and return to position. Pass thumb between the near index finger strings, pick up from below the upper near index finger string and return to position. Navaho thumb strings. Drop upper index finger loop. Draw out and turn palms slightly inward, thumb extended (Fig. 58).

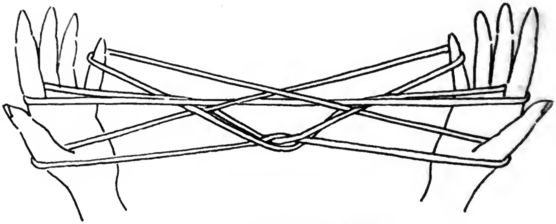


Fig. 58. Devil-devil's Anus.

XXXV. Crocodile

A continuation of Devil-devil's Anus (XXXIV)

Melville Island, Coast, East of Darwin; Wardaman, Ngainman, Djaru, Kidja, Karadjeri, Nangumarda, "Bunk"; Mardudhunera, "Spider Net"; Ngaluma, Panjima, Injibandi, Talainji, Baiong, Ingarda, "Net"; Kellerberrin, Southern Cross, "Gate"; Queensland, "Shoal of Fish."

Of the figures known at present, this pattern seems to be the most widely distributed in Australia and Oceania. Except for minor variant opening moves the procedure everywhere is the same. See D'Entrecasteaux "Number," Yap "Ten Men," New Caledonia (Nameless), New Zealand "Takapau," Fiji "Imbi," Tonga "Loukabe," Society Islands "Firi rau anini" and Marquesas "Kaukape." The final pattern also has been noted in the Caroline Islands. The Yoruba of West Africa produce a similar figure by an entirely different procedure.

1. *Both hands.* Construct 'Devil-devil's Anus" (XXXIV). Bend index finger over near index finger string and insert from below into thumb loop. Withdraw thumb from loop.

2. *Both hands.* Pass thumb under index finger strings and pick up from below on its back the near little finger string. Draw out and extend. Pass thumb between the lower and upper near index finger strings and pick up from below on its back the upper near index finger string. Draw out and Navaho thumb loops.

3. *Both hands.* Bend middle finger over the upper far index finger string, insert between strings of lower index finger loop and pick up from below the lower near index finger string. Drop little finger loop. Turn hands downward and outward working the string between index and middle fingers until figure is obtained (Fig. 59).

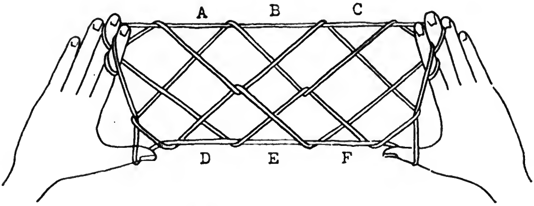


FIG. 59. Crocodile.

Release normal

XXXVI. CROCODILE'S EYE

A continuation of Crocodile (XXXV)

MELVILLE ISLAND, WARDAMAN "Vulva," NGAINMAN (not named); NGALUMA "Small Bag."

This figure also is widely distributed in Oceania but in most cases the procedure differs. A step in Yap "Carrying Money" is almost identical. Compare final pattern with Queensland "Sun Clouded Over."

1. *Both hands.* Construct Crocodile (XXXV).

2. *Both hands.* Turn hands palms downward, fingers outward. Lay figure on lap and drop all strings. Along each of the two main parallel strings which support the figure are three small triangles (Fig. 59, *A–F*).

3. *Right hand.* Move hand over figure, turn palm toward you and pick up from the far side with ball of little finger the far parallel string of the rightmost triangle (*C*). Do not lift up.

4. *Left hand.* Pick up similarly the far parallel string of the leftmost triangle (*A*).

5. *Right hand.* Bend thumb toward you, insert in rightmost triangle of the near transverse string and pick up from below the latter string (*F*). Do not lift up.

6. *Left hand.* Pick up similarly the near transverse string of the leftmost triangle (*D*).

7. *Both hands.* Lift hands with strings slack and turn palms facing by lowering elbows. Bend index finger across palm and pick up from below on its back the far thumb string. Hold secure on back of first knuckle by pressure of thumb and draw out. Turn hands fingers downward so that central figure rests on lap. Withdraw thumb and index finger from strings. Insert thumb downward into little finger loop, move hand slightly away from you and pick up from below with back of thumb the near lap string. Return to position. Pindiki and draw out (Fig. 60).

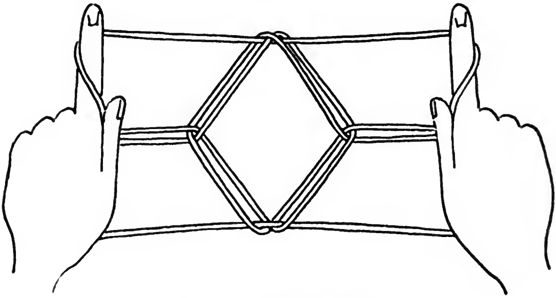


FIG. 60. Crocodile's Eye.

XXXVII. GATE

SOUTHERN CROSS, KELLERBERRIN.

This figure is the same as "Crocodile" (XXXV) but the procedure varies in the initial movements.

1. *Both hands.* Opening A.

2. *Both hands.* Drop thumb strings. Bend face over all strings, pick up with teeth the middle of the far little finger string, and draw back over all strings.

3. *Left hand.* Insert index finger from above into loop held by teeth, bend downward and pick up from below on its back the right mouth string. Draw to left.

4. *Right hand.* Pass hand to left under the left mouth string and return to right picking up this string on the right index finger. Drop loop held in mouth.

5. *Both hands.* Proceed with Movement 5 in "Devil-devil's Anus" (XXXIV) and the three movements in "Crocodile" (XXXV) (Fig. 59). (At Southern Cross and Kellerberrin Navahoing apparently always is done with the aid of the teeth.)

STRING FIGURES FORMED FROM OPENING B

XXXVIII. MEN COMING TO A FIGHT

MARAURA, WARDAMAN (not named).

This figure is almost identical in procedure with New Caledonia "Sardines" and Kiwi Papuans "A Path." In the latter instance the number of men can be diminished until none is left. This feature was not observed in either of the Australian groups. Compare finished figure with "Canoe" (Fig. 72) and "Bandicoot" (Fig. 89).

1. *Both hands.* Opening B.

2. *Both hands.* Pass thumb under index finger strings. Bend index finger down, place tip against tip of thumb and allow its loop to slip on to thumb. With little finger pick up from below the lower far thumb string and extend. Pindiki and draw out. Turn hands palms outward, fingers upward.

3. *Both hands.* Hold far index finger string secure by pressure from the middle finger. Pass thumb under the far little finger string and upward into index finger loop. With back of thumb pick up the far index finger string, withdraw index finger

and draw out with thumb on far side of little finger strings. Hold strings taut, raise elbows and turn hands fingers downward, palms toward body at level of hips. Hold near little finger string secure by pressure from middle finger. Rotate thumb upward toward body to twist its loop and at the same

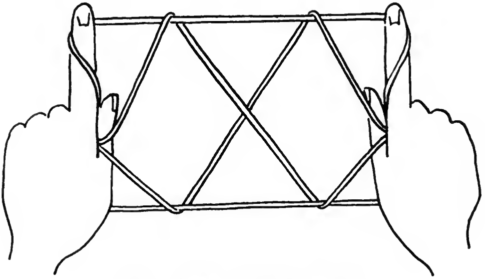


FIG. 61. Men Coming to a Fight.

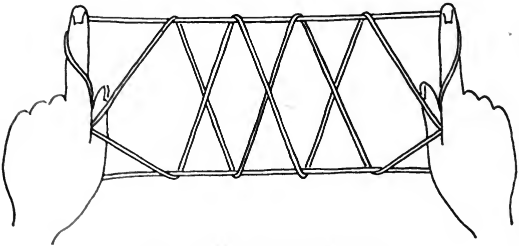


FIG. 62. Men Coming to a Fight.

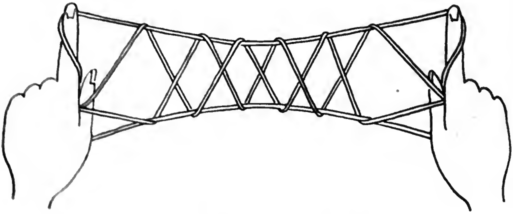


FIG. 63. Men Coming to a Fight.

time lower elbows and bring fingers upward, palms still facing body. Return to position.

4. *Both hands.* Pick up from below on back of thumb the near little finger string. Pindiki and extend (Fig: 61).

To make two more men come to the fight:

5. *Both hands.* Repeat Movements 3 and 4 (Fig. 62).

To make more men come to the fight:

6. *Both hands.* Repeat Movements 3 and 4, etc. (Fig. 63).

Release normal

XXXIX. GOANNA

WARDAMAN, NGAINMAN, INJIBANDI, MARDUDHUNERA "Spider Net," INGARDA "Net," WESTERN VICTORIA (not named).

This figure is widely distributed in Oceania but the procedures differ. Compare "Bed" (XXIV) and "Bunk" (XL).

1. *Both hands.* Opening B.

2. *Both hands.* Bend thumb down into index finger loop and pick up from below the far index finger string. Draw out and extend. Pick up from below with back of little finger the near index finger string and extend. Drop index finger loop by pushing strings off finger with tips of thumb and middle finger. Pindiki and draw out.

3. *Both hands.* Hold far index finger string secure by pressure from middle finger. (The Mardudhunera drop thumb strings and extend.) Slacken other strings and drop thumb loop. Pass thumb under all strings and pick up from below on its back the far little finger string. Draw under all strings and return to position. Place tip of thumb against tip of index finger, raise elbows slightly and allow index finger loop to slip on to thumb. Extend.

4. *Both hands.* Hold strings slack, palms facing, fingers outward. Drop little finger loop, bend little finger forward across palm and pick up from below on its back the far thumb string. Pindiki but do not extend.

5. *Right hand.* Turn hand palm downward and lower to level of waist, then turn palm inward to face body, fingers to the left, and raise to level of chin.

6. *Left hand.* At the same time as Movement 5, bring hand palm downward to level of chin, then outward, downward and

back to body at level of waist (Fig. 64). The illustration gives the point of view of the manipulator looking down on his hands.

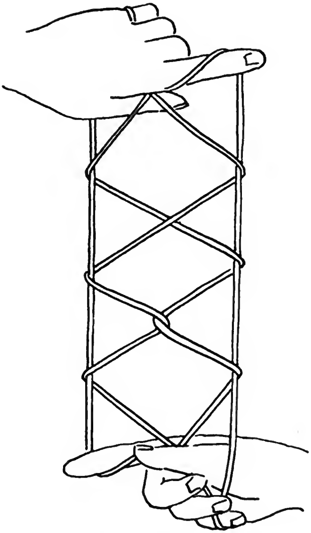


FIG. 64. Goanna.

This position of the hands can be changed to the normal extension by the following manipulation observed among the Ingarda and other tribes along the Gascoyne and Lyons Rivers.

7. *Left hand.* Move hand outward and leftward.

8. *Right hand.* Move hand downward until back of hand and central figure rest on lap. Carefully withdraw fingers from all strings so that the latter remain on lap as in Fig. 65. With fin-

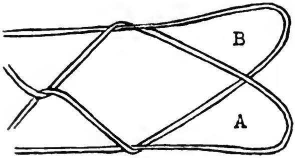


FIG. 65. A Step in "Goanna."

gers downward and palm to left insert little finger into (*A*) the near loop (formerly the index finger loop) and index finger into (*B*) the far loop (formerly the little finger loop). Extend and return to position. Insert thumb from below between little finger strings. Bend down tip of index finger, place against tip of thumb and allow index finger loop to slip on to thumb. Pindiki and extend (Fig. 66).

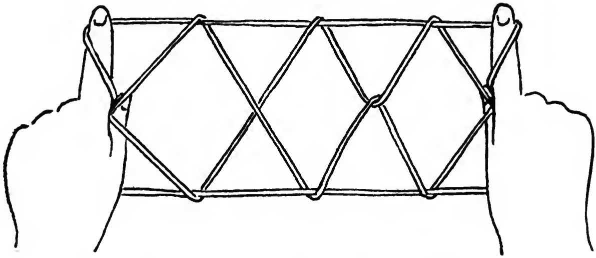


FIG. 66. Variant Extension of "Goanna" (Fig. 64).

9. *Both hands.* If figure is distorted, drop index finger strings and extend, then Pindiki and extend.

XL. BUNK

DJARU, KIDJA.

The procedure differs from "Bed" (XXIV) and "Goanna" (XXXIX). Two persons are required for this figure.

1. *First Person. Both hands.* Opening B. Face Second Person, palms outward, fingers upward.

2. *Second Person. Both hands.* Pass hands, palms downward, toward First Person under his index finger strings and over his thumb strings. Move hands slightly together. Bend little finger between First Person's thumb strings and pick up on its back his far thumb string. Turn hands outward to change string to palmar side of little finger and draw out slightly until hands are free of other strings. Pick up from below on back of thumb your palmar little finger string and lift up slightly. Pass little finger upward under the First Person's near little finger string, hook it securely and draw out under string resting on back of thumb. Turn hands palms inward, then downward

and outward until thumbs point downward. Move hands under strings of First Person and pick up from above with back of thumb his far thumb string. Return to position by lowering elbows and remove strings from hands of First Person. Pindiki but do not extend.

3. *Second Person. Right hand.* Turn hand palm downward and lower to level of waist, then turn palm inward to face body, fingers to the left, and raise to level of chin.

4. *Second Person. Left hand.* At the same time as Movement 3, bring hand with palm downward up to level of chin, then move it outward, downward and back to body at level of waist (Fig. 64). The illustration gives the point of view of the manipulator looking down on his hands.

XLI. PYTHON

WARDAMAN.

This figure is widely distributed in the Pacific Islands (and elsewhere in the world) but the procedures differ greatly. Although the manipulations are different for D'Entrecasteaux "Yam" and New Caledonia "Bracelets" it is interesting to note that these figures are made on the lap as in this North Australian example.

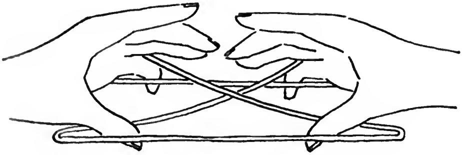


FIG. 67. A Step in "Python."

1. *Both hands.* Opening B.

2. *Both hands.* Turn hands palms downward and lay strings across knees. Drop all strings. Pass hands from the sides centerward between the two loops until they almost meet at the center. Reach fingers down between the diagonal strings and pass outward under the far diagonal string. With the ball of the little finger pick up from above the far transverse string and draw back toward you between the diagonal strings. Reach thumb back toward you over the near diagonal string and pick

up from below on its back the near transverse string (Fig. 67). Raise hands palms facing to catch strings and return to position.

3. *Both hands.* Pindiki and draw out. Turn hands palms downward and lay figure across lap. Withdraw thumb and index finger from their loops. Move hands slightly together, sliding little fingers along the far little finger string (the string nearest to you). Straighten the little finger by passing it outward under the two diagonal strings which run outward from the middle of the near transverse string, then bend it down over the far transverse string and hook it up from the far side. Bend thumb back toward you and insert under the near transverse string from the far side (Fig. 68). Return to position with the picked up strings by turning palms toward you and lowering elbows. Pindiki and draw out (Fig. 69).

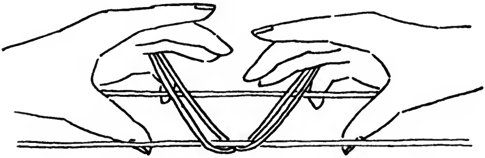


FIG. 68. A Step in "Python."

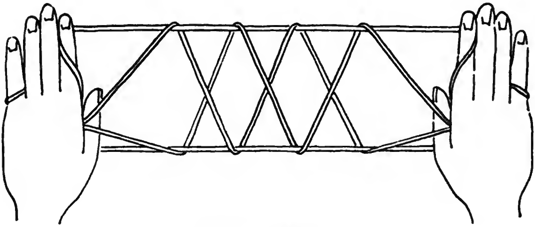


FIG. 69. Python.

Release normal

XLII. BOOMERANG

WARDAMAN.

This simple figure is identical in pattern with Queensland "Boomerang" and New Guinea "A Fish Head," the procedures of which are not given.

1. *Both hands.* Opening A.

2. *Both hands.* Bend thumb down into index finger loop and pick up from below the far index finger string. Draw out and extend. Pick up from below with back of little finger the near index finger string and extend. Drop index finger loop by pushing strings off finger with tips of thumb and middle finger.

3. *Both hands.* Pindiki and extend (Fig. 70).

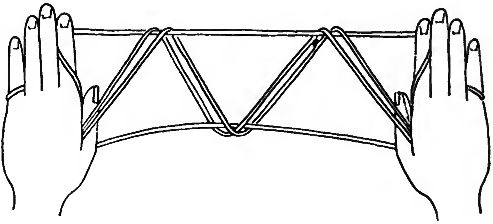


FIG. 70. Boomerang.

XLIII. CANOE

TAGOMAN, NGAINMAN, WARDAMAN, KIDJA, DJARU.

This figure is a continuation of "Boomerang" (XLII). An identical pattern is Queensland "River."

1. *Both hands.* Construct "Boomerang" (XLII).

2. *Both hands.* Drop index finger loop and with back of little finger pick up from below the far thumb string. Turn hands with fingers pointing downward and lay figure on lap. Withdraw fingers from all strings.

3. *Both hands.* With ball of little finger pick up from the far side the diagonal string. Pass thumb downward on near side of diagonal string, move it back toward you and pick up on

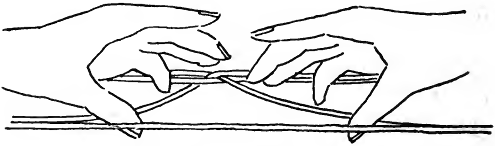


FIG. 71. A Step in "Canoe."

its back the near transverse string (Fig. 71). Lift up these strings and return to position.

4. *Both hands.* Pick up from below with the back of thumb the near little finger string near base of little finger. Extend hands, Pindiki and draw out (Fig. 72).

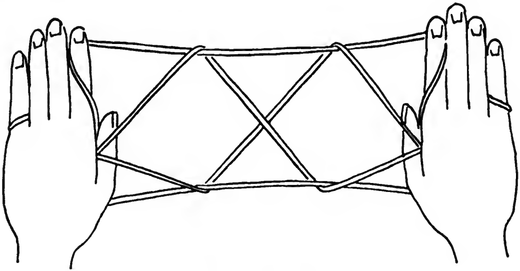


Fig. 72. Canoe.

Release normal

XLIV. Water

A continuation of Boomerang (XLII)

Wardaman.

1. *Both hands.* Construct "Boomerang" (XLII).

2. *Both hands.* Hold far index finger string secure by pressure from the middle finger. Turn hands palms outward, fingers upward. Pass thumb outward between the far index finger and far little finger strings and move hands together so that tips of

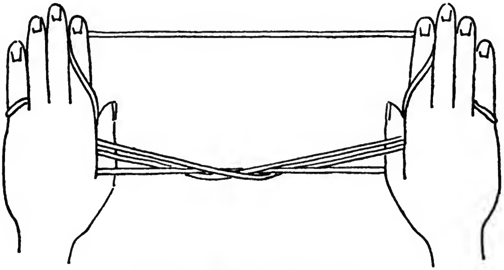


Fig. 73. Water (A River Bed).

thumb meet on the far side of the "boomerang." Bend thumb toward you and pick up on its back the two diagonal strings which comprise the side of the "boomerang." Separate the thumbs and draw out carefully. Drop the index finger strings and extend. Pindiki and draw out working the strings until figure is obtained (Fig. 73).

XLV. Blackfellow Steals a Lubra

Wardaman.

Identical patterns are Arnhem Land "Dead Man" and Queensland "Large Lizard," the procedures of which are not given.

1. *Both hands.* Opening B.

2. *Both hands.* Pass thumb over near index finger string and pick up from below the far index finger string. Extend. Pass little finger over near index finger string and pick up from below on its back the far thumb string. Extend (Fig. 74).

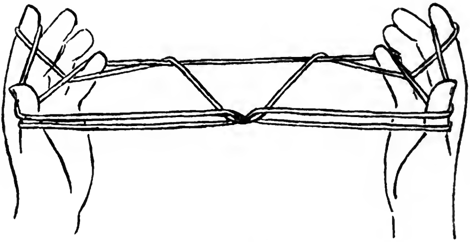


Fig. 74. A Step in "Blackfellow Steals a Lubra."

Drop index finger loop by pushing upward with tips of thumb and middle finger. Pindiki and extend (Fig. 75).

3. *Both hands.* At the very middle of the far index finger string two strings cross. Hold palms outward, fingers upward, and pick up these two strings with the teeth. Draw out several inches. Pass thumb outward and upward into little finger loop. Rotate thumb downward, then inward and upward by raising and lowering elbows to pick up the far little finger string. Drop index and little finger strings and return to position.

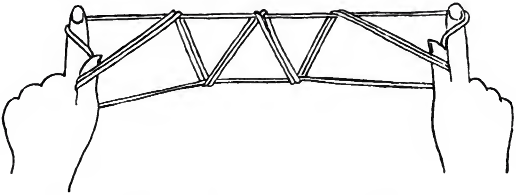


FIG. 75. A Step in "Blackfellow Steals a Lubra."

4. *Left hand.* Pass little finger from below into right thumb loop and take off.

5. *Right hand.* There are now four strings running from the teeth to the left hand. From left to right these may be termed *A, B, C* and *D.* Pass right hand from below between *B* and *C* at a point near the mouth. Move hand, palm facing you, down these strings until it reaches a loop which runs under *A,* over *B* and *C,* and under *D.* Bend the middle and index fingers away from you over this transverse string and pass middle finger under the string looped over the thumb strings, and the index

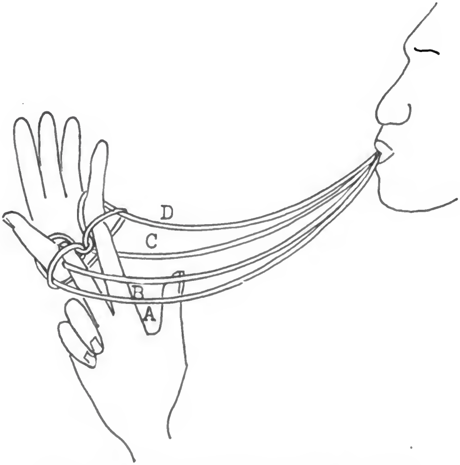


FIG. 76. A Step in "Blackfellow Steals a Lubra."

finger under the string looped over the little finger strings. Bend middle and index fingers upward to pick up the two strings indicated (Fig. 76), and draw out between *B* and *C*. Drop the strings held in the teeth. Pass the right hand to the mouth, palm toward you, and pick up with the teeth the loops on the middle and index fingers. Withdraw right hand from all strings. Turn hand palm upward, fingers outward, and insert little finger into left little finger loop from below.

6. *Left hand.* Withdraw little finger from its loop and insert from below into left thumb loop. Withdraw left thumb from its loop.

7. *Both hands.* Return to position. Reach to the mouth, pick up between the tips of thumb and index finger the strings held by teeth and remove. Pinch tips of thumb and index finger until nails meet and raise elbows so that loop slips on to index finger. Lower elbows, turn palms upward, straighten index finger and turn hands palms facing. Insert thumb from below into little finger loop and return with near little finger string and far index finger string, drawing them under the near index finger string. Drop index finger loop. Pindiki and extend (Fig. 77).

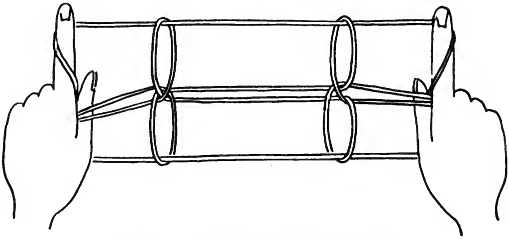


FIG. 77. Blackfellow Steals a Lubra.

Release normal

XLVI. BOY AFTER LUBRA RUNS AWAY

A continuation of Blackfellow Steals a Lubra (XLV)

WARDAMAN.

Queensland "Little Bird" has the same pattern.

1. *Both hands.* Construct "Blackfellow Steals a Lubra." The Lubra is on the left, the Blackfellow on the right.

2. *Both hands.* Turn hands palms downward, fingers outward, and lay figure across lap. Drop all strings.

3. *Right hand.* Hold hand palm downward, fingers to left, and hook with the ball of the little finger from the far side the far transverse string to the right of the "Blackfellow." Do not lift string. Bend thumb toward you and pick up from the far side the near transverse string to the right of the "Blackfellow." Do not lift string (Fig. 78).

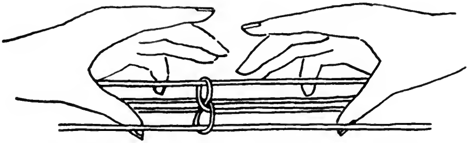


FIG. 78. A Step in "Boy After Lubra Runs Away."

4. *Left hand.* Repeat Movement 3 with left hand to pick up the same strings to the left of the "Blackfellow" (Fig. 78).

5. *Both hands.* Lift the hands with the picked up strings by lowering the elbows. Pass index finger from below into thumb loop and remove from thumb. Pass thumb under index finger strings and pick up from below the near little finger string. Insert thumb from below into index finger loop and withdraw index finger. Pindiki and extend (Fig. 79).

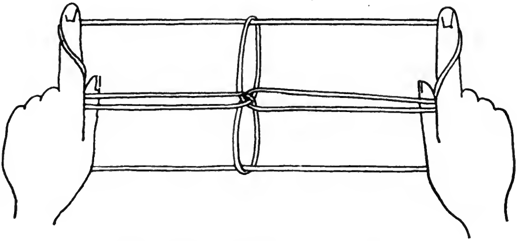


FIG. 79. Boy After Lubra Runs Away.

FLYING FOX SERIES

MELVILLE ISLAND.

XLVII. CANOE

Arnhem Land "Goanna" and "Stringray" are similar in pattern.

1. *Both hands.* Opening B.

2. *Right hand.* Pass thumb over near index finger string and pick up from below the far index finger string.

3. *Left hand.* Pass thumb under near index finger string and pick up from below the far index finger string. Draw out and extend.

4. *Both hands.* Pick up from below with back of little finger the near index finger string. Drop index finger loop by pushing upward with tips of middle finger and thumb. Pindiki but do not work strings.

5. *Both hands.* Navaho thumb strings by picking up pendant loop with teeth and lifting over thumbs. Turn hands palms upward, reach little, ring and middle fingers upward into index finger loop and bend toward you over the string just Navahoed. Press this string against palm and draw strings taut by extending thumb and index finger (Fig. 80). (Compare with Fig. 96.)

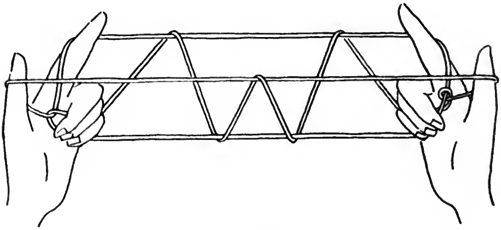


FIG. 80. Canoe.

XLVIII. FIRST FLYING FOX

A continuation of Canoe (XLVII)

The procedure is similar to New Caledonia "Bird Lime on a Tree."

6. *Both hands.* Bend thumb over middle finger. Bend index finger over thumb and allow far index finger string to slip on to

back of thumb. Extend middle, ring and little fingers, Pindiki and draw out (Fig. 81).

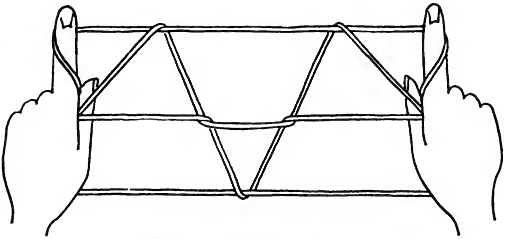


FIG. 81. First Flying Fox.

XLIX. SECOND FLYING FOX

A continuation of First Flying Fox (XLVII)

New Caledonia "Four Pigeons" is similar in procedure; South Australian "Fish Net" in pattern.

7. *Both hands.* Navaho thumb strings. Drop little finger loop. Turn hands downward and outward and extend (Fig. 82).

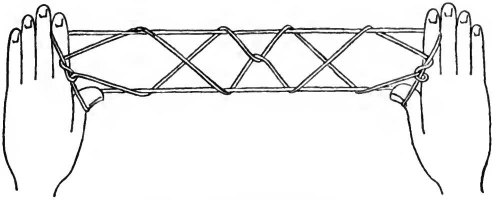


FIG. 82. Second Flying Fox.

L. THIRD FLYING FOX

A continuation of Second Flying Fox (XLIX)

8. *Both hands.* Return to position.

9. *Right hand.* With the tips of thumb and index finger grasp the intersection of strings midway between left thumb and index finger.

10. *Left hand.* Withdraw hand from all strings and pass hand palm downward over right arm. Extend little finger and

move to the left along right thumb until it can be inserted from the right into the near loop held by right thumb and index finger (previously the left thumb loop). Draw this loop slightly to the left and turn hand thumb upward. Bend thumb down and insert from below into left little finger loop. Turn thumb to right and insert into far loop held by right thumb and index finger. Return hand to position.

11. *Right hand.* Drop strings held between thumb and index finger tips.

12. *Left hand.* Repeat Movement 9 for left hand.

13. *Right hand.* Repeat Movement 10 for right hand.

14. *Left hand.* Repeat Movement 11 for left hand.

15. *Both hands.* Pindiki and extend (Fig. 83).

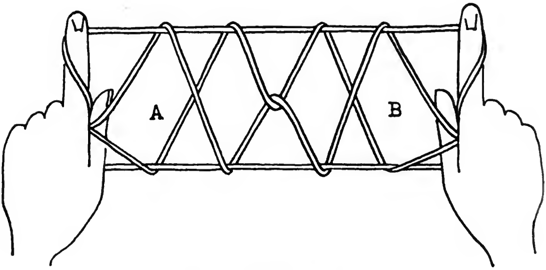


Fig. 83. Third Flying Fox.

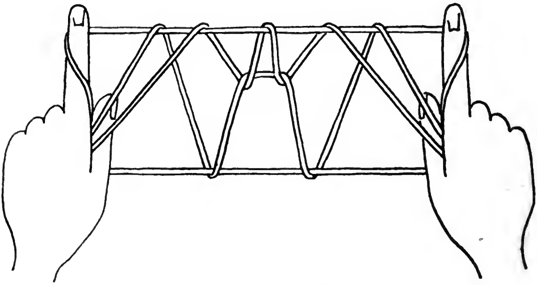


Fig. 84. Fourth Flying Fox.

LI. Fourth Flying Fox

A continuation of Third Flying Fox (L)

16. *Both hands.* Withdraw thumb from strings and pass way from you into diamond-shaped space nearest to thumb Fig. 83, *A*, *B*). Bend index finger over back of thumb and llow far index finger string to slip on to near side of thumb. Return to position. Pindiki and extend (Fig. 84).

A SERIES OF STRING FIGURES

Wardaman.

In this series of nine figures (*a–i*) two players alternate in taking a finished figure off the hands of the other and making it into a different figure. There are basic similarities with two series in the Gilbert Islands but in the latter various intermediate figures are found and several of the similar or identical figures are inverted.

LII (*a*). Boomerang

1. *First Person. Both hands.* Opening B. Construct "Boomerang" (XLII). Face Second Person, palms outward, fingers upward (Fig. 85).

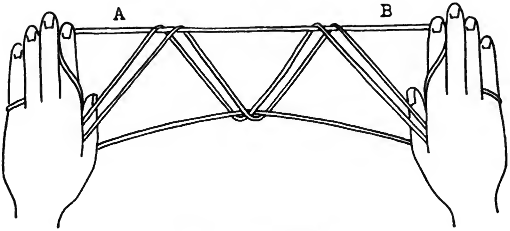


Fig. 85. Boomerang.

LIII (*b*). Python

2. *Second Person. Both hands.* With palms downward, thumbs toward you, reach over upper transverse string with little finger and pick it up close to index finger of First Person (Fig. 85, *A* and *B*). Bend thumb downward and pick up from above the lower transverse string. Turn hands upward by lowering elbows, straighten little finger, draw all strings away

from First Person and extend. With back of thumb pick up from below the near little finger string, Pindiki and extend (Fig. 86). Turn palms downward, fingers outward.

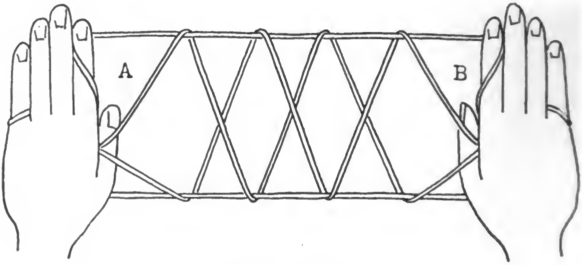


Fig. 86. Python.

LIV (*c*). Inverted Boomerang

A similar pattern is Queensland "Cassowary"

3. *First Person. Both hands.* Pass hands toward Second Person and insert thumb and little finger into triangle facing his index finger (Fig. 86, *A* and *B*). With back of thumb pick up his far index finger string and with back of little finger his near thumb string (which runs to meet the former). Withdraw strings from Second Person and extend. With back of thumb pick up from below the near little finger string, Pindiki and extend (Fig. 87). Hold palms outward, fingers upward.

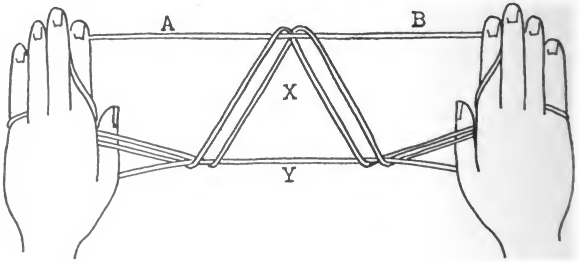


Fig. 87. Inverted Boomerang

LV (*d*). Unnamed Figure

4. *Second Person. Both hands.* With palms downward, umbs inward, pick up from above with little finger the far in-ex finger string (Fig. 87, *A* and *B*). Bend thumb downward, sert in triangle (*x*) in center of figure and pick up, on its back *y*), the lower transverse string (far little finger string). Re-ove all strings from First Person and extend. Pick up from elow on back of thumb the near little finger string, Pindiki and xtend (Fig. 88). Hold palms outward, fingers upward.

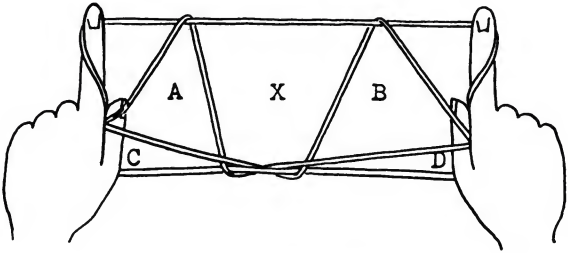


Fig. 88. Unnamed Figure.

LVI (*e*). Bandicoot

5. *First Person. Both hands.* Hold palms facing, fingers upward. The figure held by the Second Person can be divided roughly into three main parts, a large central section (Fig. 88, X), bordered on each side by a triangle. Pass the little finger into the triangle, move the hands together and pick up on the ball of the finger the diagonal string (*A*, *B*) which runs between the two hand-to-hand strings. Secure string by pressing tip of little finger against palm. Turn hands palms downward by elevating elbows and pass thumb and index finger into large central section (*X*). Enclose between thumb and index finger the string (*C*, *D*) which runs from Second Person's thumb to the lower transverse string. Lower elbows and remove all strings from Second Person. Pick up from below with back of thumb the near little finger string, Pindiki and extend (Fig. 89). Turn palms downward.

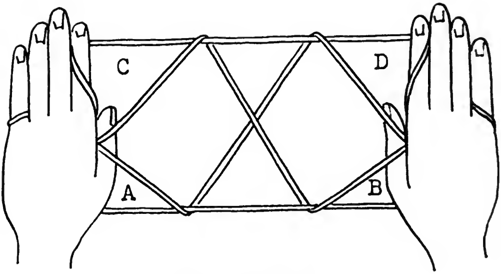


FIG. 89. Bandicoot.

LVII (*f*). UNNAMED FIGURE

6. *Second Person. Both hands.* Hold palms facing, fingers outward. Pass hands over figure and pick up from below with little finger the string which runs from back of thumb to the First Person's far little finger string (the lower transverse string, Fig. 89, *A, B*). Encircle this string with little finger and draw toward you until past all strings. Turn palms downward. Move hands away again, reach down with thumb and index on far side of your little finger string and pick up the string which runs from Second Person's thumb to his far index finger string (the upper transverse string, Fig. 89, *C, D*). Lower elbows, straighten fingers, lift all strings from First Person and return to position. Pick up from below with back of thumb the near little finger string, Pindiki and extend (Fig. 90). Hold palms outward, fingers upward.

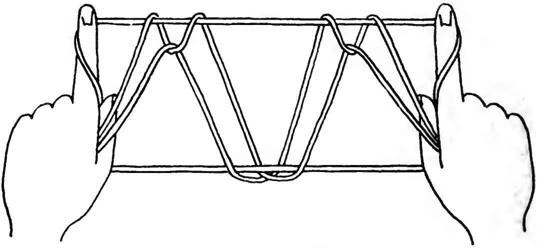


FIG. 90. Unnamed Figure.

LVIII (*g*). UNNAMED FIGURE

7. *First Person.* Repeat Movement 2 (Fig. 91). Hold palms downward.

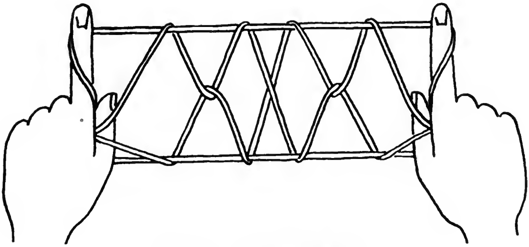


FIG. 91. Unnamed Figure.

LIX (*h*). UNNAMED FIGURE

8. *Second Person.* Repeat Movement 3 (Fig. 92). Hold palms outward.

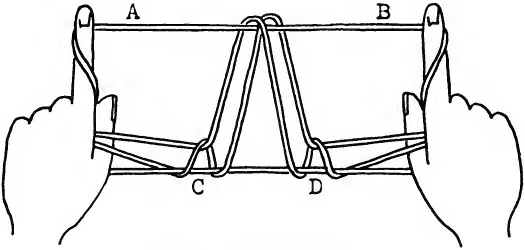


FIG. 92. Unnamed Figure.

LX (*i*). WATER (A RIVER BED)

9. *First Person. Both hands.* Hold palms outward, thumbs downward. Pick up from below with little finger the upper transverse string close to Second Person's index finger (Fig. 92, *A, B*). Bend thumb downward and insert into small triangle which faces the lower transverse string about one-third toward center from hand (Fig. 92, *C, D*). Pick up the transverse string on back of thumb by lowering elbows, lift all strings from Second

Person and extend. With back of thumb pick up from below the near little finger string, Pindiki and extend (Fig. 93).

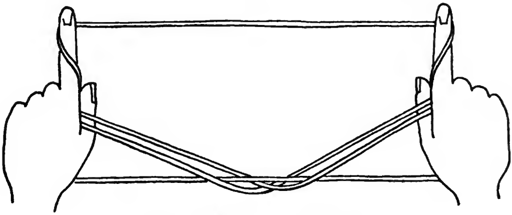


FIG. 93. Water (A River Bed).

STRING FIGURES FORMED FROM OPENING C

LXI. FOUR BLACKFELLOWS

NGAINMAN.

An identical pattern is Queensland "Four Boys Walking in a Row." A fairly long string is required for this figure.

1. *Both hands.* Opening C.

2. *Both hands.* With thumb pick up from below the upper near index finger string and draw out. With little finger pick up from below the upper far index finger string and draw out. Drop upper loop from index finger and extend.

3. *Right hand.* With middle finger pick up from below the left palmar string and return.

4. *Left hand.* With middle finger pick up the right palmar string between the middle finger strings and extend.

5. *Right hand.* Hold strings secure, pass hand over left hand and with the tips of thumb and index finger pick up the left far little finger string. Carry this string over the back of the left little and ring fingers and drop between the left ring and middle finger (thereby changing the little finger loop to a ring finger loop in inverted position).

6. *Left hand.* Repeat Movement 5 with left hand. Return to position.

7. *Right hand.* Turn hand palm toward you and bend little, ring and middle fingers over all strings to hold them secure. Roll hand to left by raising elbow and with the index finger and thumb lift up the left far thumb string, the index and middle finger strings and the near ring finger string.

8. *Left hand.* Bend the little finger forward over the far ring finger string and under the strings held up by the right hand and pick up from below the left near thumb string. Draw out and return to position.

9. *Left hand.* Repeat Movement 7 with left hand.

10. *Right hand.* Repeat Movement 8 with right hand.

11. *Both hands.* Drop thumb loop. Bend thumb over near index finger, pass it away from you under all other strings and return with the far little finger string on its back.

12. *Both hands.* The near index finger string now runs under the thumb strings, thence up over the near thumb string to the center of the figure. Turn hands palms toward you. Bend little finger down over middle finger strings and pick up from below on its back the aforementioned string at a point midway between its intersection with the near thumb string and the center of the figure (Fig. 94). Return to position and Navaho little finger loops.

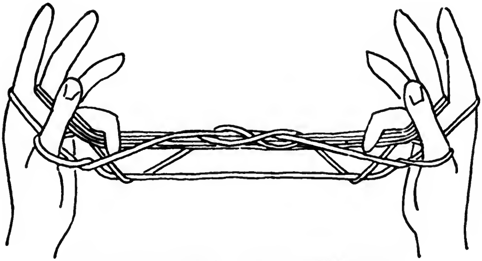


FIG. 94. A Step in "Four Blackfellows."

13. *Right hand.* Reach index, middle and ring fingers over back of left hand and pull the loops on the left index, middle and ring fingers over the ends of the fingers so that they drop on the palmar side.

14. *Left hand.* Repeat Movement 13 with the left hand.

15. *Both hands.* Draw the hands slightly apart until the loops dropped are free. Bend index finger over far thumb string and pick up from below the near thumb string. Straighten the finger and withdraw the thumb to transfer thumb loop to the index finger in an inverted position. Pass thumb under index finger strings and pick up from below the near

little finger string. Pass thumb from below into index fing loop and withdraw index finger. Extend slightly, Pindiki a draw out, working the strings slowly until figure is obtain (Fig. 95).

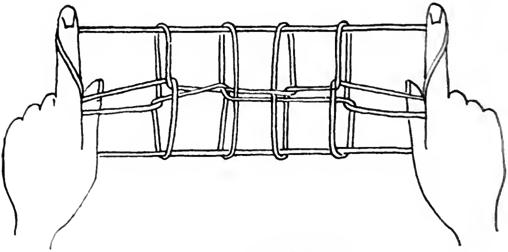


FIG. 95. Four Blackfellows.

Release normal

LXII. CANOE

WARDAMAN.

This figure has the same name and procedure in Queensland.

1. *Both hands.* Opening C.

2. *Right hand.* Pass thumb between upper and lower near index finger strings and pick up from below the far upper index finger string. Return to position.

3. *Left hand.* Pass thumb between upper and lower near index finger strings and pick up from below the far lower index finger string. Return to position.

4. *Both hands.* Pass ring finger over the upper near index finger string, pick up from below the lower near index finger string and extend. Hold taut the far ring finger string and move the thumbs slightly together. This movement should cause the upper near index finger string, which runs from hand to hand, to swing beneath the other strings.

5. *Both hands.* Turn hands so that fingers point upward, palms outward, and let the long hanging string swing to the level of the thumbs. Pass index finger downward into thumb loop, then move it away from you over the hanging string. Hook the index finger over this hanging string and return to position with it by rotating the finger and lowering the elbows.

urn hands palms downward, fingers outward. Drop thumb trings and extend.

6. *Both hands.* Pass thumb downward between ring finger trings and return with the near index finger string which runs rom hand to hand (Fig. 96). (Compare with Fig. 80.)

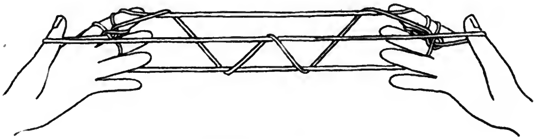


FIG. 96. Canoe.

LXIII. NET

SWAN REACH.

This figure is found in North Australia, Central Australia and Queensland under various names.

1. *Both hands.* Opening C.

2. *Right hand.* Pass thumb between upper and lower near index finger strings and pick up from below the far upper index finger string.

3. *Left hand.* Pass thumb between the upper and lower near index finger strings and pick up from below the far lower index finger string.

4. *Right hand.* Pass middle, ring and little fingers toward you over the far lower index finger string and press tips of fingers against palm of hand. Straighten the middle finger and pick up from below the near lower index finger string.

5. *Left hand.* Pass middle, ring and little fingers toward you over the far upper index finger string and press tips of fingers against palm of hand. Straighten the middle finger and pick up from below the lower near index finger string.

6. *Both hands.* Keep little and ring fingers pressed against palm and turn hands palms toward you so that upper near index finger string assumes a palmar position across the middle finger. Bend the middle finger over this string and rotate the finger downward, outward and upward to wind this string once around the first joint. This movement causes the former middle finger loop to slip off the finger. Extend.

7. *Both hands.* The string just dropped by the middle finger now runs from a point where it passes over the near middle finger string to the near side of the index finger. Pick up this string from the far side with the back of thumb. Navaho near thumb strings. Return to position and drop little finger strings.

8. *Both hands.* Turn hands palms downward. Bend the thumbs toward each other, thence downward and outward to wind the near thumb string once around the thumb. Extend and work strings until figure is obtained (Fig. 97).

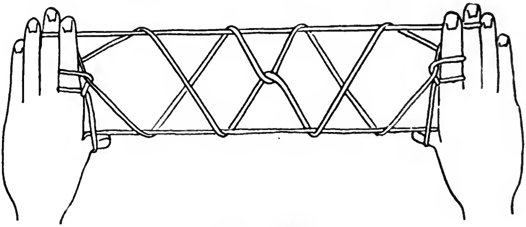


FIG. 97. Net.

LXIV. TURTLE

NGARLA, NGALUMA "Spider Net."

Arnhem Land "Tortoise" is similar. Two persons and a long string are required.

1. *First Person. Both hands.* Opening C.

2. *First Person. Both hands.* Treat the double loop as a single string and execute again Opening C to place a quadruple loop on the index fingers.

3. *Second Person. Left hand.* Insert index finger from above into loops held by First Person near his right hand and remove from his right index finger.

4. *Second Person. Right hand.* With palm downward pick up from above with ball of ring finger one of the near strings a few inches from left index finger of First Person and draw out slightly. Pass index finger over this string, pick up from above with its ball a second string and draw out over first string.

5. *Second Person. Right hand.* Straighten index finger until the back of the first joint rests under ring finger string. Move thumb along ball of index finger to catch palmar index finger string on back of first joint of thumb. Separate thumb and in-

dex finger slightly. Pass index finger forward and pick up from above with its ball a third string. Draw this string out over second string.

6. *Second Person. Right hand.* Proceed similarly to Movement 5 and pick up the fourth near string and each of the four far strings of the quadruple loop. When the eighth string has been drawn out on the ball of index finger pass the index finger outward under all strings toward right hand of First Person.

7. *First Person. Right hand.* Pass index finger downward into loop on right index finger of Second Person and take off.

8. *Second Person. Right hand.* Pass index finger downward into ring finger loop and take off.

9. *Both Persons. Both hands.* Extend hands slightly to draw strings taut (Fig. 98).

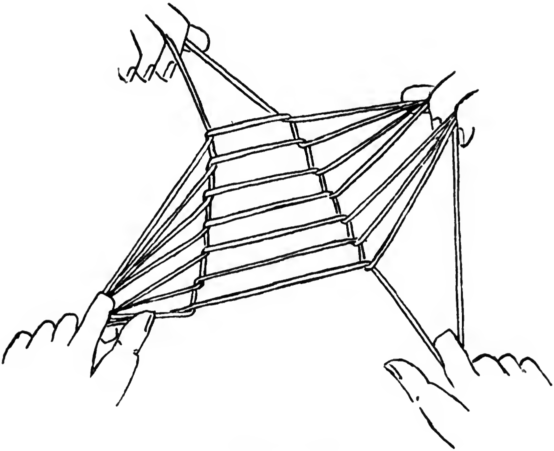


Fig. 98. Turtle.

STRING FIGURES FORMED FROM ELBOW OPENING

LXV. Tomahawk

Wagoman.

1. Elbow Opening.

2. *Right hand.* With tips of thumb and index finger pick up the near left arm string, carry over back of left hand and place

between left ring and little fingers. Drop string from right hand. Turn hand palm upward, fingers outward. Move hand away from you and pick up between the little and ring fingers the string suspended from the far side of the left ring finger.

3. *Both hands.* Draw hands apart and assume usual position but be sure that the left arm loop is secure above the left elbow.

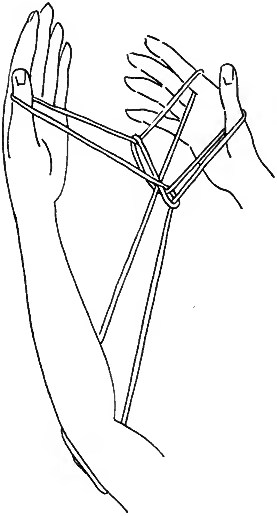


Fig. 99. Tomahawk.

Turn hands palms downward so that far ring finger string runs across palms. Bend thumb down under this string from the near side and, by turning hands palms facing, pick up on back of thumb. Return to position.

4. *Left hand.* Proceed as in Opening A by picking up the right palmar string from below with the index finger (the right palmar string must be picked up first).

5. *Right hand.* With the index finger pick up from below the left palmar string and draw out between the left index finger strings.

6. *Right hand.* Bend thumb over the far thumb string, pass under the index finger strings and the little finger strings and return with the far little finger string. Draw up through the original thumb loop and allow the latter to slip off the thumb and on to the new loop (Navaho). Drop little finger loop.

7. *Left hand.* Drop index finger loop. Drop little finger loop and extend (Fig. 99).

LXVI. Lightning

Wagoman.

1. Elbow Opening.

2. *Right hand.* Pass hand between body and elbow loop, pick up from near side with the fingers, palms upward, the far left arm string about half way to bottom of loop and draw up near body.

3. *Left hand.* Pick up from the far side with the fingers, palms upward, the near left arm string about halfway to bottom of loop.

4. *Both hands.* Bring hands toward each other and turn palms downward by elevating elbows. Slightly overlap the index fingers, right hand outmost, to form a small pendant loop. Insert in this loop from the far side the index, middle and ring fingers and lock the crossed strings by bringing the tips of these fingers against tip of thumb. Hold strings secure and point fingers downward so that left arm loop will fall over left hand and on to strings. This movement will also cause the far ring finger string to hang below the hands in its original position of a pendant loop.

5. *Both hands.* Turn hands palms facing and raise fingers slowly until they point outward thus changing the aforementioned far ring finger string to a far little finger string. Separate thumb and fingers and extend carefully. There should be two loops on each thumb and one on each little finger. Pindiki.

6. *Right hand.* Turn hand palm downward, fingers outward.

7. *Left hand.* Close little, ring and middle fingers against palm. Rotate hand outward and downward, then inward and upward to wind once on the index finger all the left hand strings (Fig. 100).

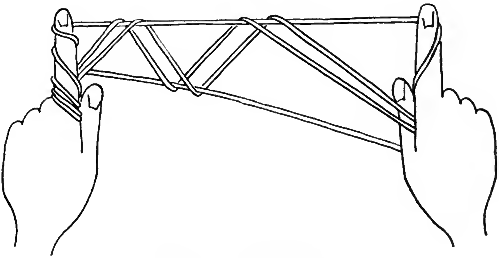


FIG. 100. Lightning.

LXVII. AXE

WARDAMAN.

1. Elbow Opening.

2. *Right hand.* Slide elbow loop upward over left shoulder. Pass hand between body and loop and pick up from below between tips of thumb and index finger, palm upward, the far (back) shoulder string about half way to bottom of loop. Draw up near body.

3. *Left hand.* Pick up from below between tips of thumb and index finger, palm upward, the near (front) shoulder string about half way to bottom of loop.

4. *Both hands.* Bring hands toward each other and turn palms downward by elevating elbows. Slightly overlap the index fingers, right hand outmost, to form a small pendant loop. Insert index finger in this loop from the far side and lock the crossed strings by locking tip against tip of thumb. Extend hands carefully. Straighten index finger and raise thumb to place a loop on each thumb and index finger. Shrug shoulder to cause shoulder loop to slide over left hand on to strings. Extend and return to position.

5. *Both hands.* Pass thumb over near index finger string and pick up from below the far index finger string. Bend little finger over near index finger string and pick up from below the far thumb string. Return to position and drop index finger loop by pushing strings from below with tips of thumb and middle finger.

6. *Both hands.* Press far thumb strings against index finger and Pindiki.

7. *Right hand.* Turn hand palm downward.

8. *Left hand.* Close little, ring and middle fingers against palm. Rotate hand downward and outward, then inward and upward to wind once on the index finger the left hand strings (Fig. 101).

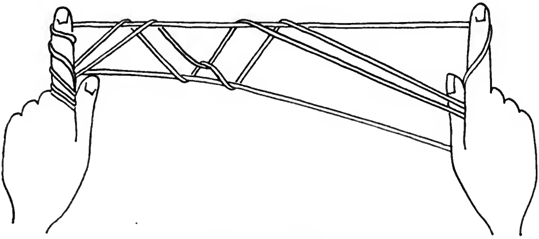


FIG. 101. Axe.

STRING FIGURES FORMED FROM ODD OPENINGS

LXVIII. HOUSE

WARDAMAN.

Two strings and two persons are required for this figure.

This figure is widely distributed in the Pacific Islands. With only trivial differences the procedure is similar in Queensland "Turtle," Yap "House," New Zealand "Moutohora," Marquesas and Hawaii. Comparisons should be made with similar figures from Queensland, Marin-anim of New Guinea, and Palau, the manipulations of which are not given.

1. *Both Persons.* Sit on ground face to face with legs crossed. Pick up First String, form a square and lay on ground between the four feet. Pick up Second String and make another square by inserting hands from below to form wrist loops.

2. *First Person. Both hands.* With thumb and index finger lift up the Second Person's near wrist string and draw toward you.

3. *Second Person. Both hands.* Turn hands downward by raising elbows, pass hands away from you under the string held up by First Person, and pick with thumb and index finger his near wrist string.

4. *Both Persons. Both hands.* Draw hands apart and permit wrist loops to slip over hands on to strings by raising el-

bows. With ball of index finger hook up corner of square made by First String, draw up through loop, and permit latter to slip off finger and on to strings (Fig. 102). Place new loop over foot.

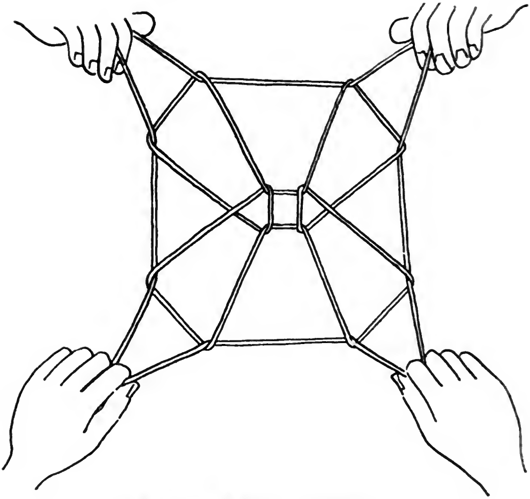


FIG. 102. A Step in "House."

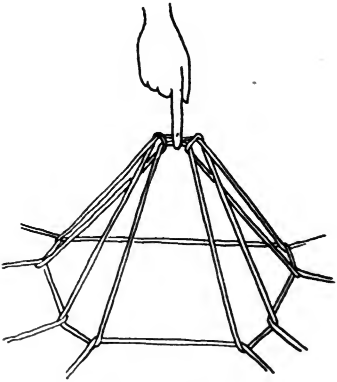


FIG. 103. "House."

5. *First Person. Right hand.* Reach into center of figure and lift up between thumb and index finger the two central strings (Fig. 103).

STRING TRICKS

LXIX. "Cutting Off the Head"

Wardaman, Southern Cross (not named by either).

1. *Right hand.* Place string over head so that a long loop hangs in front of body.

2. *Both hands.* With thumb and index finger pick up middle of string of pendant loop. Raise hands to level of mouth and cross them with strings tightly drawn. Pick up with teeth the intersection of crossed strings and bring hand back to a position directly in front of shoulder. Extend index finger so that string becomes a thumb loop (Fig. 104). Swing hands upward and backward to toss far thumb string over head to nape of neck. Retain loop on thumb and return hand to shoulder position, palms outward, fingers upward. Bring hands suddenly together with a loud clap and at the same time release strings from teeth. Both strings now appear on back of neck.

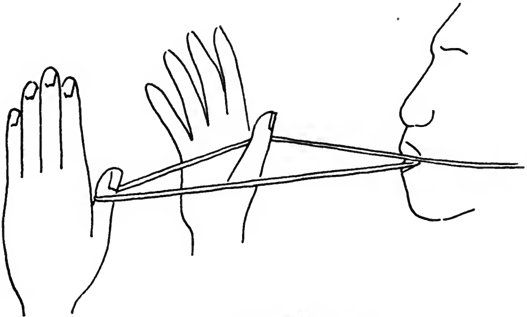


Fig. 104. "Cutting Off the Head."

LXX. "Hanging"

Tagoman, Wardaman, Brinken, Mullik-Mullik, Swan Reach, Southern Cross (not named by these tribes).

1. *Right hand.* Place string over head so that a long loop hangs in front. Pick up right loop string, carry upward to left

to wind once tightly around neck, and drop to form a pendant loop again.

2. *Right hand.* Pass hand over right string of loop, pick up middle of left loop string between tips of thumb and index finger and return to position.

3. *Left hand.* Pick up middle of right loop string with tips of thumb and index finger and return to position.

4. *Both hands.* Bring hands together so that the strings cross near index finger. Encircle the two strings with index finger and lock by pressure from tip of thumb. Pass middle, ring and little fingers into loop which hangs below. Draw hands apart and turn palms facing (Fig. 105). Raise hands, insert head into loop from below and drop all strings.

5. *Either hand.* Pull any string and all strings will come off neck.

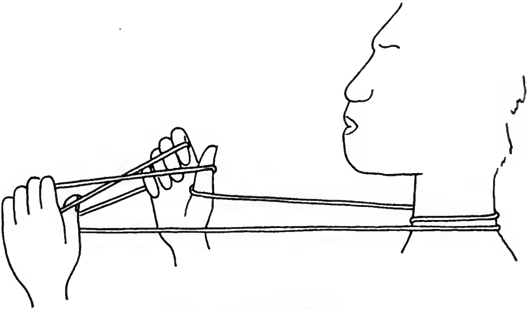


Fig. 105. "Hanging."

LXXI. Bullocky

Wardaman, Southern Cross (not named), Ngaluma "Fingers Caught."

1. *Left hand.* Hold hand with fingers and thumb outward.

2. *Right hand.* Hang loop over left thumb and hand and withdraw right hand. Insert index finger from the left under near thumb string, reach between thumb and index finger and pick up from the right with ball of finger the far thumb string. Draw out to the left of near thumb string to form a small loop, rotate index finger inward and upward to give loop a twist, and place it over left index finger.

3. *Right hand.* In a similar manner pick up the far hand string between the index and middle fingers and place, after twisting, on the left middle finger.

4. *Right hand.* Proceed similarly by drawing out and twisting a loop from between the middle and ring fingers and placing on the left ring finger, and from between the ring and little fingers and placing on the little finger (Fig. 106).

5. *Right hand.* Draw taut the two long hanging strings. Lift left thumb loop off thumb and drop. Now pull on the near long hanging string and all the strings will come off the fingers.

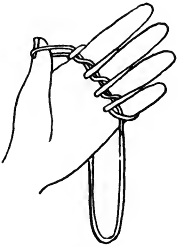


FIG. 106. Bullocky.

LXXII. "CUTTING OFF THE FINGERS"

DARWIN, MOGUMBER, SWAN REACH (not named in these places).

1. *Right hand.* Place string across left palm, palm upward, fingers outward, so that long loop hangs below. With thumb and index finger grasp both strings of hanging loop and draw them upward in such a manner that the near string passes between little and ring fingers and the far string between index and middle fingers. Draw strings taut, move hand to left so that strings pass between left index finger and thumb, then move hand back to right to carry strings over left thumb (Fig. 107).

2. *Left hand.* Bend index finger over upper thumb near string and by rotating the finger downward and outward pick up a loop on this finger.

3. *Right hand.* Draw lower near left thumb string between left little and ring fingers and drop all strings from right hand (Fig. 108). With thumb and index finger pick up the two left

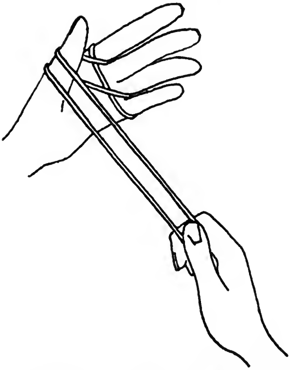


FIG. 107. A Step in "Cutting Off the Fingers."

thumb loops, draw tightly between the left middle and ring fingers and drop on far side (Fig. 109). Draw out the left palmar string and all the strings will be free.

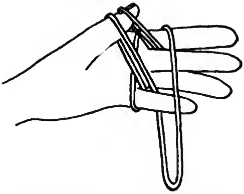


FIG. 108. A Step in "Cutting Off the Fingers."

LXXIII. CUTTING OFF THE NOSE

NGALUMA.

This trick is the same as Torres Strait "A Fly on the Nose," except that the loop is made in reverse order.

1. *Both hands.* Hold string between tips of thumb and index finger, with palms downward and hands about four inches apart, so that a long loop hangs below.

2. *Right hand.* Circle hand forward, then back to the left until tips of thumbs meet to form a small loop.

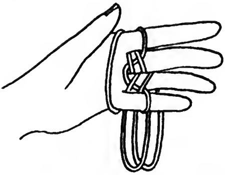


FIG. 109. A Step in "Cutting Off the Fingers."

3. *Both hands.* Grasp crossed strings between tips of thumb and index finger and place in teeth so that the small loop hangs from the mouth. Drop all strings from hands.

4. *Left hand.* Grasp both strings of the long hanging loop and hold taut with strings uncrossed.

5. *Right hand.* Pass hand forward through long loop above the left hand. Turn palm upward and raise hand until index finger is directly in front of small loop held by teeth. Bend index finger down into small loop (Fig. 110), hook finger over lower part of a loop and rotate finger leftward, then upward to give loop a twist (Fig. 111). Place tip of index finger against tip of nose, palm facing you.

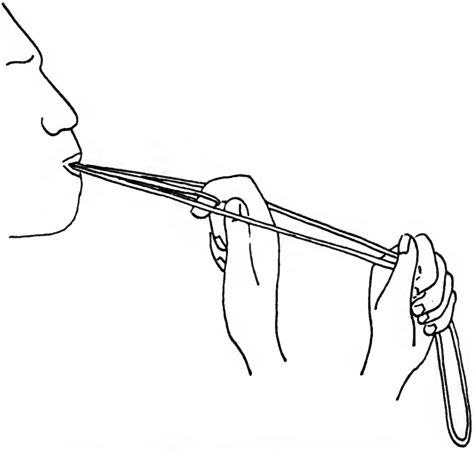


FIG. 110. A Step in "Cutting Off the Nose."

6. *Left hand.* To "cut off the nose" give the long loop a jerk and at the same time release the string from the teeth.

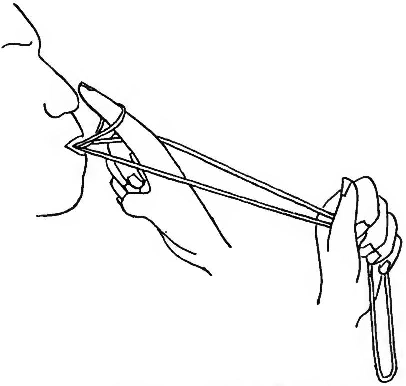


FIG. 111. A Step in "Cutting Off the Nose."

LXXIV. "THREADING THE NEEDLE"

NGALUMA, MOGUMBER, SWAN REACH (not named in these areas).

A single piece of string is better for this widely distributed catch, although a loop used as a single string will do.

1. *Right hand.* Place string over left thumb, palm inward, fingers to right, so that about ten inches of string hangs on the far side of thumb. With right thumb and index finger pick up the near thumb string and wind several times around thumb.

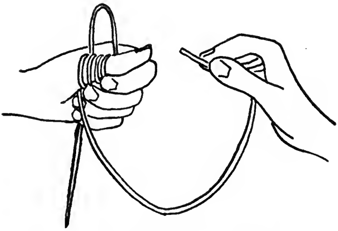


FIG. 112. "Threading the Needle."

Make a small loop in the remainder of string held in right hand and place this between left thumb and index finger so that the long hanging part of the string is near the base of the left thumb.

2. *Right hand.* With tips of right thumb and index finger pick up the hanging end of the original far thumb string and draw to right without pulling taut. Hold this string as if you were threading a needle and make several passes toward the loop (Fig. 112). Finally draw this string tight and allow it to pass upward between the left thumb and index finger at the same time as you pass the right hand quickly over the left hand. This movement draws the string up into the loop but gives the impression that it has been passed through the loop.

APPENDIX A

String Figures from Arnhem Land, North Australia

These figures were collected by Mr. H. U. Shepherdson at Milingimbi, North Australia, and mounted on paper for the South Australian Museum. The manipulations were not recorded but since many of the patterns are the same as those collected by the author a few hundred miles to the southwest it seems safe to conclude that some of the procedures in the two districts are identical.

There is one local peculiarity not noted in any other region, the presence of several twists in the strings at various points of the patterns. This arrangement presumably is a distinctive contribution to string figure construction.

1. *Stingray Spear.*—This figure belongs to the same class as Lily (Fig. 25), Headdress (Fig. 27), Hut (Fig. 28), Fish Spear (Fig. 40) and Spear (Fig. 42). There are four loops of which two are on the same finger.

2. *Goanna.*—Identical with Bunk (Fig. 44), Men Coming to a Fight (Fig. 61), Canoe (Fig. 72) and Bandicoot (Fig. 89).

3. *Stingray.*—The same as the preceding with added twists in the strings which run from each diamond to the thumb.

4. *Crocodile.*—The same as Net (Fig. 97), Second Flying Fox (Fig. 82) and Third Flying Fox (Fig. 83) except for the twists in the strings which run from each lateral diamond to the thumb.

5. *Wallaby's Intestines.*—Similar to the preceding except that the twists separate the two central diamonds.

6. *Sun.*—The same as Crocodile's Eye (Fig. 60).

7. *Dead Man.*—The same as Blackfellow Steals a Lubra (Fig. 77).

8. *Tortoise.*—Similar in type to Dilly Bag (Fig. 50), Turtle (Fig. 51) and Turtle (Fig. 98).

9. *Mutchichi.*—Identical with Dilly Bag (Fig. 17). A doubled string is employed.

10. *Honey Comb.*—The same as Bed (Fig. 43) and Goanna (Figs. 64 and 66).

11. *Crab.*

12. *Native Companion.*

13. *Fire Sticks.*

14. *Mopoak.*

15. *Fresh Water Turtle.*

16. *Turtle Shell.*—The same as South Queensland Turtle.

APPENDIX B

String Figures from Hermannsburg, Central Australia

These figures, mounted on paper in the South Australian Museum, were collected by Mr. N. B. Tindale and others at Hermannsburg, Central Australia, from mixed Arunta and Kukatja natives. The procedures were not recorded. In many specimens the strings have slipped, hence comparison is difficult. However, many patterns can be identified as identical with North Australian designs. It will be seen that several patterns have a variety of names.

At the present time there seem to be no outstanding features to distinguish Central Australian figures from those of other areas.

1. *Deep Water Hole with Ripples.*—Similar to Devil-devil's Anus (Fig. 58).

2. *Woman.*—
3. *Cloud.*
4. *Lizard.*
5. *Frog.*
6. *Woman with Child.*
7. *Dogs Copulating.*

These figures are the same as Bunk (Fig. 44), Men Coming to a Fight (Fig. 61), Canoe (Fig. 72) and Bandicoot (Fig. 89).

8. *Cooked Rabbit on a Stick.*—Of the same general type as Lily (Fig. 25), Headdress (Fig. 27), Hut (Fig. 28), Fish Spear (Fig. 40) and Spear (Fig. 42).

9. *Rabbit Bandicoot.*—Suggestive of Emu Foot (Fig. 29).

10. *Two Dogs.*—Similar to the preceding except that the design is at each end.

11. *Testicles.*—
12. *Stone.*
13. *Crayfish.*
14. *Native House.*

Except for slight alterations in the extensions these four figures seem to be the same.

15. *Large Lizard.*—The same as a step in Blackfellow Steals a Lubra (Fig. 76).

16. *Baby Evacuating.*—The same as Fourth Flying Fox (Fig. 84).

17. *Wooden Baby Carrier.*—Identical with Opening A (Fig. 3).

18. *Little Bird.*—Resembles Blackfellow (Fig. 22) and Boy after Lubra Runs Away (Fig. 79).

19. *Beetle.*—
20. *Fishing Net.*
21. *Carpet Snake.*
22. *Bird.*
23. *Snake.*
24. *Eagle.*
25. *Little Dog.*
26. *Centipede.*

These figures are comparable with Men Coming to a Fight (Fig. 62) and Python (Figs. 69 and 86).

27. *Mythical Large Watersnake.*—The same as Goanna (Figs. 64 and 66) and Bed (Fig. 43).

28. *Gently Running Water.*—Apparently the same as Crocodile (Fig. 59).

29. *Man Sitting Down.*—The same as First Flying Fox (Fig. 81).

30. *Man* (Pattern of Seat in Sand).
31. *Penis.*
32. *Little Boy.*
33. *Rock Hole.*
34. *Bull Ant.*
35. *Four Men Sitting Down.*
36. *Two Women Fighting.*
37. *Bed.*
38. *Ngappa Crow Walking with Wings Drooped.*
39. *Butterfly.*
40. *Duck.*
41. *Two Women Sitting Down.*
42. *Ring Necked Parrot.*

43. *Ground Lark.*
44. *Bandicoot.*
45. *Monkey.*
46. *Chickens.*
47. *Small Fish.*
48. *Butterfly.*
49. *Big Spider.*
50. *Camel Rope.*
51. *Eaglehawk.*
52. *Kivilakmila.*
53. *Two Kangaroos.*
54. *Native Mice Playing.*
55. *Yertitjina.*
56. *Dog Chasing an Emu.*
57. *Bird.*
58. *Witchety Grub Mob.*
59. *Long Edible Snake.*
60. *Two Intertwined Snakes.*
61. *Two Arunta Natives.*
62. *Quarrel Between Two Women with Sticks.*
63. *Creek in Flood.*
64. *Dingo Pups.*
65. *Four Food Bowls.*
66. *Kangaroo.*
67. *Ants.*
68. *Marsupials.*
69. *Two Arunta Women Hitting Each Other.*
70. *Platform of Branches on which Churinga are Laid.*
71. *Yards for Mustering Stock* (in imitation of fish weirs).
72. *Hairy Caterpillar Found in Eucalyptus Tree.*
73. *Pregnant Dingo.*
74. *Pouched Rat.*
75. *Winter Flood.*
76. *Waterholes Dug in Sand of a Creek in Dry Country.*
77. *Lizard.*
78. *Lizard.*

APPENDIX C

String Figures from the Coorong, South Australia

This figure was collected by Mr. N. B. Tindale of the South Australian Museum where it is now mounted on a card. The manipulation was not recorded. In view of Cawthorne's report of string figures among the Adelaide tribe in 1844 it seems likely that other figures were formerly known in the nearby Coorong district.

1. *Fish Net.*—The design is identical with Second Flying Fox (Fig. 82), Third Flying Fox (Fig. 83) and Net (Fig. 97).

APPENDIX D

STRING FIGURES FROM THE DIAMANTINA DISTRICT, SOUTH AUSTRALIA

These figures, mounted on paper in the South Australian Museum, were collected by Mr. N. B. Tindale in the Diamantina district in the northeastern part of the state. The manipulation was not recorded.

1. *Yam.*—This figure belongs to the same general class as Lily (Fig. 25), Headdress (Fig. 27), Hut (Fig. 28), Fish Spear (Fig. 40) and Spear (Fig. 42). There are four separate loops.

2. *Man Stalking Emu.*—There are three progressive figures in a series.

APPENDIX E

Comparative Chart of Australian String Figure Distributions

CAPITALS—Same or very similar procedure.
Italics—Same figure, procedure not given.
Parentheses—Same figure, different procedure.

Figs.	North Australia	Western Australia	Central Australia	Queensland	Victoria, South Australia	Torres Strait	D'Entre-Casteaux Is. [and New Guinea]	Yap, Nauru, Caroline Is., Gilbert Is.	New Caledonia	Fiji and Polynesia	Elsewhere
I 9	TWO DOGS Wardaman Ngainman										
II 10	A WATERHOLE Wardaman Ngainman									(One Eye [1]) Hawaii	
III 11	TWO WATERHOLES Wardaman Ngainman			SNAKE [2] *Deaf Adder* [3]		A FIGURE [4]		FLINT AND STEEL [5] Yap		UTOSUVI [6] Fiji	
IV 12–13	DEAD MAN IN A TREE Wagoman										
V 14	PHALLUS Wardaman										
VI 15–17	DILLY BAG Wardaman *Mutchichi* Arnhem Land			Compare *Dilly Bag* [7]							
VII 18–19	BABY BEING BORN Wardaman										
VIII 20–21	BIG RIVER BAT Wardaman										
IX 22	BLACKFELLOW Melville Is. Coast, East of Darwin PENIS Ngainman			Compare (Giant Crane) (Flying Fox) [7a]			A HISS [8]	ONE CHIEF [9] Yap Compare *Palau* [10] and *Nauru* [11] figures	LOOPER CATERPILLAR [12]	BANUVE [13] Fiji	
X 23–26	LILY Wardaman Ngainman STINGRAY SPEAR Arnhem Land Compare with XI and XXI	Compare with XI and XXI	Compare with XI and XXI	LILY ROOT [14] Compare with XI and XXI	YAM Diamantina, South Australia					BROOM [15] Hawaii	

XI 27	Compare with X and XXI	HEADDRESS Ngaluma Compare with X and XXI	COOKED RABBIT ON A STICK Compare with X and XXI	Compare with X and XXI							
XII 28		YAM Djaru Kidja HUT Nangumarda Panjima Injibandi Talainji Ngaluma Baiong Targudi Kellerberrin Southern Cross PENIS Ingarda									
XIII 29		EMU FOOT Gnowangerup Quairading	*Rabbit-Bandicoot* *Two Dogs*								
XIV 30		EMU Southern Cross Laverton									
XV 31	TWO KANGAROOS Wardaman Ngainman						*A Figure*[16] [Marind-anim] A step in (Seine[17])	"NO NAME"[18] Yap KANI MUMUM[19] Gilbert Is.		VAHIA[20] Society Is.	
XVI 32					TWO TURTLES ON A LOG Maraura TWO SWANS[21]						Compare (Carrying Wood[22]) Navaho
XVII 33	KANGAROO TRAIL Wardaman										
XVIII 34–37	DEAD KANGAROOS Wardaman						(Seine[23])	(Te Kabanei[24]) Gilbert Is.			

---

[1] Dickey, p. 19.
[2] Stanley, p. 82; Fig. 8, p. 77.
[3] Roth, Pl. 6, Fig. 3.
[4] Haddon, A. C., p. 17.
[5] Jayne, p. 320.
[6] Hornell, p. 28. Minor variations.
[7] Roth, Pl. 12, Fig. 6.
[7a] Roth, Pl. 5, Fig. 5; Pl. 4, Fig. 3.
[8] Jenness, p. 316. Procedure same as Yap. See fn. 9.
[9] Jayne, p. 253. Procedure the same as North Australia except for opening on thumb instead of little finger.
[10] Raymund, Pl. 6, No. 1.
[11] Hambruch, p. 348; Pl. 17, No. 4.
[12] Compton, p. 228. See fn. 9.
[13] Hornell, p. 19. See fn. 9.
[14] Stanley, p. 76; Fig. 1, p. 77. Minor variations.
[15] Dickey, p. 141. Minor variations.
[16] Wirz, Pl. 24, No. 3.
[17] Jenness, p. 309. Different procedure.
[18] Jayne, p. 176. Minor variations.
[19] Maude, p. 33. Kani Mumun has variant extension. Kan Raure has different procedure.
[20] Handy, p. 73. Identical except for intermediate steps to form Te Tuma O Te Papa.
[21] Haddon, K., 1918, p. 127.
[22] Haddon, A. C., 1912, p. 46. Different procedure.
[23] Jenness, p. 309. Different procedure.
[24] Maude, p. 47. Different procedure.

APPENDIX E—*Continued*

		North Australia	Western Australia	Central Australia	Queensland	Victoria, South Australia	Torres Strait	D'Entre-Casteaux Is. [and New Guinea]	Yap, Nauru, Caroline Is., Gilbert Is.	New Caledonia	Fiji and Polynesia	Elsewhere
XIX	38	Big Mob of Blackfellows Wardaman			*Rain* *Two Tomahawks*[25]							
XX	39	Armband Wardaman							Three Stars[26] Yap			
XXI	40	Fish Spear Wardaman Compare with X and XI	Compare with X and XI	Compare with X and XI	Compare with X and XI		Compare Fish Spear[27]				Compare Balawa "B"[28] Fiji	Cf. Panjo India Hornell, 1932
XXII	41	Vulva Wardaman										
XXIII	42			Spear[29]								
XXIV	43	Compare with XXXIX and XL	Bed Southern Cross Gnowangerup Moora	*Mythical Large Watersnake*	Compare with XXXIX and XL	Compare with XXXIX and XL						
XXV	44		Bunk Djaru Kidja Naugumarda	*Woman* *Cloud* *Lizard* *Woman with Child* *Dogs Copulating*								
XXVI	45	Eaglehawk Wardaman										
XXVII	46–49	Big Pig Wardaman										Compare Parrot Cage[30] Yoruba, West Africa
XXVIII	50	Dilly Bag Wardaman *Tortoise* Arnhem Land			Compare *Tortoise*[31]				Compare (Ri Ni Kaokao[31a]) Gilbert Is.			Compare with XXVII
XXIX	51	Turtle Wagoman Compare Tortoise Arnhem Land	Turtle Djaru Kidja		Compare with XXVIII							

XXX	52		Turtle Djaru Kidja									
XXXI	55	Two Black-fellows Wardaman										
XXXII	56	Bamboo Wardaman Ngainman							House of the Blos Bird [32] Yap		Te Faro No Oro [33] Society Is.	
XXXIII	57	Rain Wardaman Ngainman										
XXXIV	58	Devil-Devil's Anus Ngainman See XXXV	A step in XXXV		Mouth of a Fish [34]			A step in XXXV	A step in XXXV	A step in XXXV	A step in XXXV Vasua [35] Fiji	
XXXV	59	Continuation of XXXIV Crocodile Melville Is. Coast, East of Darwin Bunk Wardaman Ngainman A step in XXXVI	See XXXIV Bunk Djaru Kidja Nangumarda Karadjeri Spider Web Mardudhunera Net Panjima Injibandi Talainji Ngaluma Ingarda Baiong Gate Kellerberrin Southern Cross A step in XXXVI		Continuation of XXXIV Shoal of Fish [36] See XXXVI			See XXXIV Number [37]	See XXXIV Ten Men [38] Yap (A Figure [39]) Caroline Is. A step in XXXVI	See XXXIV "Nameless" [40]	See XXXIV Takapau [41] New Zealand Imbi [42] Fiji Loukabe [43] Tonga Firi Rau Anini [44] Society Is. Kaukape [45] Marquesas Compare with XXXVI	Compare [46] (A Figure) Yoruba, West Africa (Owl's Nest) Klamath, Oregon

---

[25] Roth, Pl. 11, Figs. 3, 8.

[26] Jayne, p. 167.

[27] Rivers and Haddon, p. 149, Fig. 1; Haddon, A. C., 1912, p. 323.

[28] Hornell, p. 13.

[29] Secured at Swan Reach, South Australia, from a visiting Arunta.

[30] Haddon, A. C., 1912, p. 33.

[31] Roth, Pl. 7, Fig. 2.

[31a] Maude, p. 42.

[32] Jayne, p. 162. Procedure inverted.

[33] Handy, p. 91. Minor variations.

[34] Stanley, p. 80; Fig. 5, p. 77. Reversed procedure and different order of opening.

[35] Hornell, p. 43. Identical procedure.

[36] Stanley, p. 80; Fig. 6, p. 77. Reversed procedure.

[37] Jenness, p. 306. Variant opening moves.

[38] Jayne, p. 150. Variant opening moves.

[39] Raymund, p. 47.

[40] Compton, p. 218. Opening made with left hand.

[41] Andersen, p. 25. Variant opening moves.

[42] Hornell, p. 45.

[43] Hornell, p. 67.

[44] Handy, p. 67.

[45] Handy, p. 29.

[46] The Yoruba and Klamath procedures are entirely different. See Jayne, p. 69.

	North Australia	Western Australia	Central Australia	Queensland	Victoria, South Australia	Torres Strait	D'Entre-Casteaux Is. [and New Guinea]	Yap, Nauru, Caroline Is., Gilbert Is.	New Caledonia	Fiji and Polynesia	Elsewhere
XXXVI 60	Continuation of XXXV Crocodile Eye Melville Is. Vulva Wardaman "Not Named" Ngainman *Sun* Arnhem Land	Continuation of XXXV Small Bag Ngaluma		See XXXV *Sun Clouded over*[47]				Continuation of XXXV A step in Carrying Money[48] Yap *A Figure*[49] Palau (Taai; Boua-n[50]) Gilbert Is. (The Sun[51]) Gilbert Is.	(Butterfly[52])	(Wahine[53]) New Zealand (Tavola[54]) Fiji (Huma[55]) Tonga. A continuation of XXXV (Punawai o Makena[56]) Hawaii (Hono Vahine[57]) Society Is. A continuation of XXXV (I Puava Akani[58]) Marquesas. A continuation of XXXV	
XXXVII 59		Gate Kellerberrin Southern Cross									
XXXVIII 61–63	"Not Named" Wardaman *Goanna* Arnhem Land *Stingray* Arnhem Land Has twists		Figure 61 *Woman* *Cloud* *Lizard* *Frog* *Woman with Child* *Dogs Copulating* Figure 62 *Beetle* *Fishing Net* *Carpet Snake* *Bird* *Eagle* *Little Dog* *Centipede* *Dingo Pups*		Men Coming to a Fight Maraura		[A Path[59]] Kiwi Papuans		Sardines[60]		
XXXIX 64–66	Goanna Wardaman Ngainman *Honey Comb* Arnhem Land Compare with XXIV and XI.	Goanna Injibandi Spider Web Mardudhunera Net Ingarda Compare with XXIV and XI.	*Mythical Large Water Snake*	*Honey*[61] (Frog[62])	A Figure Western Victoria			Compare (Turtle[63]) Yap (A Figure[64]) Tokelau and Gilbert Is.		Compare (Figures) New Zealand[65] Fiji[66] Tonga[67] Samoa[68] Society Is. and Marquesas[69]	Compare (Bogabo Diamonds[70]) Philippines

XL 64	Compare with XXIV and XXXIX	**Bunk** Djaru Kidja									
XLI 67–69	**Python**[71] Compare with XXXVIII, Fig. 62		Compare with XXXVIII, Fig. 62	(Shrimp[72]) *Four Shrimps*[73]			(Yam[74])	(Caroline Islands Diamonds[75]) (Na Ubwebwe[76]) Gilbert Is. *Figure*[77] Nauru *Figure*[78] Ellice, Gilbert and Tokelau Is.	(Bracelets[79])	(Figure) New Zealand[80] Marquesas[81] Fiji[82]	
XLII 70	**Boomerang** Wardaman See XLIII	A step in XLIII		*Boomerang*[83]			[*A Fish Head*[84]]	(Na Umake[85]) Gilbert Is. Compare also (Kan Tie) in a series (A Figure[86]) Gilbert and Ellice Is.			
XLIII 71–72	Continuation of XLII **Canoe** Wardaman Ngainman Tagoman Compare with XXXVIII and XLI See XLIV	Continuation of XLII **Canoe** Djaru Kidja	Compare with XXXVIII and XLI	*River*[87]				Compare with (Two Diamonds[88]) New Zealand			

---

[47] Roth, Pl. 10, Fig. 1.
[48] Jayne, p. 160.
[49] Raymund, p. 55.
[50] Maude, pps. 27, 50. Different procedure.
[51] Grimble, p. 214. Different procedure.
[52] Compton, p. 214. Different procedure.
[53] Andersen, p. 20. Different procedure.
[54] Hornell, p. 66. Different procedure.
[55] Hornell, p. 67. Derived from Loukabe (No. XXXV) by entirely different procedure.
[56] Dickey, p. 56. Different procedure.
[57] Handy, p. 69. Derived from Firi rau anini (No. XXXV) by entirely different procedure.
[58] Handy, p. 31. Derived from Kaukape (No. XXXV) by entirely different procedure.
[59] Haddon, K., 1930, p. 78. Similar procedure.
[60] Compton, p. 224.
[61] Roth, Pl. 8, Fig. 3.
[62] Stanley, p. 79. Different procedure.
[63] Jayne, p. 265.
[64] Hornell, p. 79. Similar to Fiji.
[65] Andersen, p. 35.
[66] Hornell, p. 51. Different procedure.
[67] Hornell, p. 66. Different procedure
[68] Hornell, p. 75. Different procedure.
[69] Handy, pps. 44, 86. Different procedure.
[70] Jayne, p. 43. Same as New Zealand. See fn. 64.
[71] Executed on lap as in New Guinea and New Caledonia but with different procedure.
[72] Stanley, p. 79. Different procedure.
[73] Roth, Pl. 8, Fig. 2.
[74] Jenness, p. 305. Executed on lap as in North Australia and New Caledonia but with different procedure.
[75] Jayne, p. 260. Similar to New Zealand.
[76] Maude, p. 12. Different procedure.
[77] Hambruch, p. 348, Fig. 3.
[78] Hornell, p. 77. Similar to Caroline Islands.
[79] Compton, p. 211. Executed i lap. Same as New Guinea.
[80] Andersen, p. 89. Different pro cedure.
[81] Handy, p. 42. Different pro cedure.
[82] Hornell, p. 34. Different pro cedure.
[83] Roth, Pl. 9, Fig. 7.
[84] Holmes, p. 286.
[85] Maude, pps. 16, 21.
[86] Hornell, p. 77. Somewhat dif ferent.
[87] Roth, Pl. 11, No. 4.
[88] Andersen, p. 91.

APPENDIX E—*Continued*

		North Australia	Western Australia	Central Australia	Queensland	Victoria, South Australia	Torres Strait	D'Entre-Casteaux Is. [and New Guinea]	Yap, Nauru, Caroline Is., Gilbert Is.	New Caledonia	Fiji and Polynesia	Elsewh
XLIV	73	Continuation of XLIII WATER										
XLV	74–77	BLACKFELLOW STEALS A LUBRA Wardaman *Dead Man* Arnhem Land See XLVI		*Large Lizard* Fig. 76					(Two Chiefs[89]) Yap			
XLVI	78–79	Continuation of XLV BOY AFTER LUBRA RUNS AWAY Wardaman		*Little Bird*								
XLVII	80	Flying Fox Series CANOE Melville Is. See XLVIII *Goanna* Arnhem Land *Stingray*[90] Arnhem Land										
XLVIII	81	Continuation of XLVII FIRST FLYING FOX Melville Is. See XLIX			Compare (Seagull Flying Upside Down[91])					BIRDLIME ON A TREE[92]		
XLIX	82	Continuation of XLVIII **SECOND FLYING FOX** Melville Is. **See L** ***Crocodile*** Arnhem Land ***Wallaby's Intestines***[90] Arnhem Land				*Fish Net* Coorong, South Australia				FOUR PIGEONS[93]		

L	83	Continuation of XLIX THIRD FLYING FOX Melville Is. See LI										
LI	84	Continuation of L FOURTH FLYING FOX Melville Is.		*Baby Evacuating*	Compare (Seagull Flying Right Side Up[94])							
LII	85	A Series of Nine Figures BOOMERANG Wardaman See LIII										
LIII	86	Continuation of LII PYTHON Wardaman See LIV							(Two series have similar figures[95]) Gilbert Is.		Compare (Pareru[96]) New Zealand (Resembles a series in Fiji)	
LIV	87	Continuation of LIII INVERTED BOOMERANG Wardaman See LV			*Cassowary*[97]	(Water Rat[98]) South Australia						
LV	88	Continuation of LIV UNNAMED Wardaman See LVI										
LVI	89	Continuation of LV BANDICOOT Wardaman See LVII										

---

[89] Jayne, p. 188. Different procedure.

[90] With added twists, a local peculiarity.

[91] Stanley, p. 85. Different procedure.

[92] Compton, p. 227. Very similar.

[93] Compton, p. 227. Similar.

[94] Stanley, p. 85. Different procedure.

[95] Maude, pp. 12ff, gives two series which include most of these figures but with various intermediate steps. Different procedure.

[96] Andersen, p. 94.

[97] Roth, Pl. 4, Fig. 7.

[98] Haddon, K., 1918, p. 126.

APPENDIX E—*Continued*

		North Australia	Western Australia	Central Australia	Queensland	Victoria, South Australia	Torres Strait	D'Entre-Casteaux Is. [and New Guinea]	Yap, Nauru, Caroline Is., Gilbert Is.	New Caledonia	Fiji and Polynesia	Elsewhere
LVII	90	Continuation of LVI UNNAMED Wardaman See LVIII										
LVIII	91	Continuation of LVII UNNAMED Wardaman See LIX										
LIX	92	Continuation of LVIII UNNAMED Wardaman See LX										
LX	93	Continuation of LIX WATER Wardaman										
LXI	94–95	FOUR BLACK-FELLOWS Ngainman			*Four Boys Walking in a Row*[99]							
LXII	96	CANOE Wardaman			*Canoe*[100] (Canoe[101]) CANOE[102]							
LXIII	97	*Crocodile*[90] Arnhem Land *Wallaby's Intestines*[90] Arnhem Land		*Marsupials*	*Emu*[103] *Four Pronged Spear*[104] Compare (Flying Foxes[105])	NET Swan Reach *Fish Net* Coorong, South Australia	(Flying Foxes[106])				Compare (Figures) New Zealand[107] Samoa[108] Fiji[109]	Widely distributed[110]
LXIV	98	*Tortoise* Arnhem Land	TURTLE Ngarla Ngaluma									
LXV	99	TOMAHAWK Wagoman							ELBOW OPENING Palau[110a]			

LXVI	100	Lightning Wagoman							See LXV			
LXVII	101	Axe Wardaman							See LXV			
LXVIII	102–103	House Wardaman			Turtle[111] Compare (Canoe[112]) (Fish net[113])			Compare [A Figure[114]] Marind-anim	House[115] Yap *A Figure*[116] Palau		A Figure New Zealand[117] Marquesas[118] Hawaii[119]	
LXIX	104	“Cutting Off the Head” Wardaman	Not Named Southern Cross Ingarda								New Zealand[120]	Negritoes[121] Philippines
LXX	105	“Hanging” Wardaman Tagoman Brinken Mullik-mullik	Not Named Southern Cross Ingarda			Not Named Swan Reach, South Australia			Hanging[122] Yap			
LXXI	106	Bullocky Wardaman	Not Named Ingarda Southern Cross Fingers Caught Ngaluma				Not Named[123]		Mouse[124] Yap	Uprooting Mania[125]		Philippines, Africa, North America[126]
LXXII	107–109	“Cutting Off the Fingers” Darwin	Not Named Mogumber Ingarda			Not Named Swan Reach, South Australia					New Zealand[127] Marquesas[128] Hawaii[129]	Germany, Scotland[130]
LXXIII	110–111		“Cutting Off the nose” Ngaluma Ingarda				Fly on the Nose[131]					
LXXIV	112		Snare Ngaluma Mogumber Ingarda			Not Named Swan Reach, South Australia						Threading the Needle[132] Widely distributed

---

[99] Roth, Pl. 3, Fig. 3.

[100] Roth, Pl. 12, Fig. 2.

[101] Stanley, p. 92. Different procedure.

[102] Haddon, K., 1918, p. 131. Almost identical procedure.

[103] Roth, Pl. 4, Fig. 5.

[104] Roth, Pl. 12, Fig. 1.

[105] Haddon, K., 1918, p. 133.

[106] At Cape York, North Queensland. See preceding footnote.

[107] Andersen, p. 73. Different procedure.

[108] Hornell, p. 72.

[109] Hornell, p. 11.

[110] Haddon, A. C., 1912, p. 73. See also Cunnington, p. 125.

[110a] Raymund, No. 43, p. 50.

[111] Stanley, p. 84.

[112] Roth, Pl. 12, Fig. 4.

[113] Roth, Pl. 12, Fig. 5.

[114] Wirz, Pl. 24, Fig. 11.

[115] Jayne, p. 200. Similar.

[116] Raymund, p. 46.

[117] Andersen, p. 137.

[118] Handy, p. 48. Identical.

[119] Dickey, p. 127.

[120] Andersen, p. 147.

[121] Jayne, p. 339.

[122] Jayne, p. 339.

[123] Rivers and Haddon, p. 192.

[124] Jayne, p. 340.

[125] Compton, p. 233.

[126] Jayne, p. 320.

[127] Andersen, p. 147.

[128] Handy, p. 52.

[129] Dickey, p. 150.

[130] Haddon, A. C., 1912, p. 84

[131] Jayne, p. 348.

[132] Jayne, p. 354; Davidson, p. 393.

APPENDIX F

Cup and Saucer

This figure, learned in Philadelphia from an American-born Scotch girl of immigrant parents, is included herein as a matter of record. No information is available as to its ultimate origin. Comparison should be made with "Dressing a Skin" of the Thompson Indians, British Columbia, given by Jayne, p. 30.

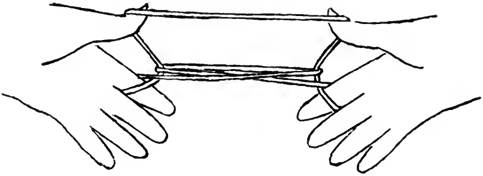


Fig. 113. Cup and Saucer.

1. *Both hands.* Opening A.

2. *Both hands.* Bend thumb over near index finger string and pick up from below on its back the far index finger string.

3. *Both hands.* Navaho thumb strings, drop little finger strings and extend (Fig. 113).

Release: Drop index finger strings and extend

BIBLIOGRAPHY

Andersen, J. C. 1927. Maori String Figures. *Board of Maori Ethnological Research.* Memoir 2.

Boas, F. 1888. The Game of Cat's Cradle. *International Archiv für Ethnographie*, **1**.

Bunce, D. 1857. Australasiatic Reminiscences. Melbourne.

Cawthorne, W. A. (1844). Rough Notes on the Manners and Customs of the Natives. Royal Geographical Society (South Australian Branch). Adelaide, 1925–1926.

Compton, R. H. 1919. String Figures from New Caledonia and the Loyality Islands. *Journal of the Royal Anthropological Institute*, **49**.

Davidson, D. S. 1927. Some String Figures of the Virginia Indians. *Indian Notes.* Museum of the American Indian (Heye Foundation), **4**, No. **4**.

1938. A Preliminary Register of Australian Tribes and Hordes. The American Philosophical Society. Philadelphia.

1938. An Ethnic Map of Australia. *Proceedings of the American Philosophical Society*, **79**, No. 3.

Dieffenbach, E. 1843. Travels in New Zealand. 2 vols. London.

Dickey, L. A. 1928. String Figures from Hawaii. *Bishop Museum Bulletin* 54.

EYRE, E. J. 1845. Journals of Expeditions into Central Australia, etc. 2 vols. London.

GILL, W .W. 1876. Life in the Southern Isles. London.

GORDON, G. B. 1906. Notes on the Western Eskimo. *Transactions of the Free Museum of Science and Art*, **2**, No. 1. Philadelphia.

GRIMBLE, A. 1931. Gilbertese Astronomy and Astronomical Observances. *Journal of the Polynesian Society*, **40**.

HADDON, A. C. 1903. A Few American String Figures and Tricks. *American Anthropologist*, **2**.

1906. String Figures from South Africa. *Journal of the Royal Anthropological Institute*, **36**.

1912. Report of the Cambridge Anthropological Expedition to Torres Straits, **4**. Cambridge.

HADDON, K. 1911. Cat's Cradles from Many Lands. London.

1918. Australian String Figures. *Proceedings of the Royal Society of Victoria*, **30**.

1930. Artists in String. London.

HALL, C. F. 1864. Life with the Esquimaus. London.

HAMBRUCH, P. 1914. Nauru. *Ergebnise der Südsee-Expedition*, 1908–1910. Part 2B, No. 1. Hamburg.

HANDY, W. C. 1925. String Figures from the Marquesas and Society Islands. *Bishop Museum Bulletin* 18.

HOLMES, J. H. 1924. In Primitive New Guinea. London.

HORNELL, J. 1927. String Figures from Fiji and Western Polynesia. *Bishop Museum Bulletin* 39.

1932. String Figures from Gujarat and Kathiawar. *Memoirs of the Asiatic Society of Bengal*, **11**: **4**.

JAYNE, C. F. 1906. String Figures. New York.

JENNESS, D. 1920. Papuan Cat's Cradles. *Journal of the Royal Anthropological Institute*, **50**.

1924. Eskimo String Figures. *Report of the Canadian Arctic Expedition*, 1913–1918, **13**, Part B. Ottawa.

KLUTSCHAT, H. W. 1881. Als Eskimo unter den Eskimos. Wien.

KROEBER, A. L. 1899. The Eskimo of Smith Sound. *American Museum of Natural History*, Bulletin 12.

MAUDE, H. C. and H. E. 1936f. String Figures from the Gilbert Islands. *Journal of the Polynesian Society*, **45f.**

RIVERS, W. H. R., and HADDON, A. C. 1902. A Method of Recording String Figures and Tricks. *Man*, No. 109.

ROTH, W. E. 1902. North Queensland Ethnography. Bulletin 4. Brisbane.

1903. Report. *American Anthropologist*, **5**.

1924. An Introductory Study of the Arts, Crafts, and Customs of the Guiana Indians, *Bureau of American Ethnology*, 38th Annual Report. Washington.

STANLEY, G. A. V. 1926. String Figures of the North Queensland Aborigines. *Queensland Geographical Journal*, **60**, **61**. Brisbane.

SPENCER, B. 1928. Wanderings in Wild Australia. 2 vols. London.

SMYTH, R. B. 1878. The Aborigines of Victoria. 2 vols. Melbourne.

TENICHEFF, W. 1898. Games of the Eskimaux. *L'Activité de l'homme*. Paris.

TYLOR, E. B. 1879. Remarks on the Geographical Distribution of Games. *Journal of the Royal Anthropological Institute*, **9**.

RAYMUND, P. 1911. Die Faden und Abnehmespiele auf Palau. *Anthropos*, **6**.

WIRZ, P. 1922. Die Marind-anim von Holländische-Süd-Neu-Guinea. Hamburg.

INDEX TO VOLUME 84

Allen, Henry Butler, Alexander Dallas Bache and his connection with The Franklin Institute of the State of Pennsylvania, 145
Anderson, Thomas F., and B. M. Duggar, The effects of heat and ultraviolet light on certain physiological properties of yeast, 661
Aurora and geomagnetism (Gartlein), 299
Australian string figures, Aboriginal (Davidson), 763

Bache, Alexander Dallas, Commemoration of the life and work of, 125
Basilica, Early Christian, S. Pietro in Vincoli and the tripartite transept in (Krautheimer), 353
Berkner, L. V., Contributions of ionospheric research to geomagnetism, 309
Binaries, eclipsing, Determination of limb darkening in, from color-index observations (Kopal), 565
Bone, Responsive (Davenport), 65
Bowring, John, and the poetry of the Slavs (Coleman), 431
Bronze, Corroded, of Corinth (Caley), 689

Caley, Earle R., The corroded bronze of Corinth, 689
Cheyney, Edward P., The connection of Alexander Dallas Bache with the University of Pennsylvania, 151
Colbert, Leo Otis, Alexander Dallas Bache as Superintendent of United States Coast Survey, 1843–1867, 173
Coleman, Arthur Prudden, John Bowring and the poetry of the Slavs, 431
Columbus, What — saw on landing in the West Indies (Olschki), 633
Conklin, Edwin G., Alexander Dallas Bache and his connection with the American Philosophical Society, 125
Cosmic radiation, Production of neutrons by (Korff), 589

Davenport, C. B., Responsive bone, 65
Davidson, Daniel Sutherland, Aboriginal Australian string figures, 763
Drosophila melanogaster, Structure and development of centrifuged eggs and early embryos of (Howland), 605
Duggar, B. M. (with Thomas F. Anderson, see Anderson)

Electrons, secondary, Production of, by electrons of energy between 0.7 and 2.6 MEV (Hornbeck and Howell), 33

Fleming, J. A., Geomagnetism: worldwide and cosmic aspects with especial reference to early research in America, 263

Gaposchkin, Cecilia Payne (with Fred L. Whipple, see Whipple)
Gartlein, C. W., Aurora and geomagnetism, 299
Geomagnetism, Symposium on, 187
Gesell, Arnold, The genesis of behavior form in fetus and infant, 471
Gish, O. H., Terrestrial electricity in relation to geomagnetism, 187

Hallborg, H. E., Correlations of short wave radio transmission across the Atlantic with magnetic conditions, 323
Heck, N. H., The magnetic survey of the United States, 205
Heyl, Paul R., Magnetism and its uses, 339
Hornbeck, George, and Irl Howell, Production of secondary electrons by electrons of energy between 0.7 and 2.6 MEV, 33
Howell, Irl (with George Hornbeck, see Hornbeck)
Howland, Ruth B., Structure and development of centrifuged eggs and early embryos of *Drosophila melanogaster*, 605
Hyla crucifer, Reproductive physiology of (Rugh), 617

Ionospheric research, Contributions of, to geomagnetism (Berkner), 309

Jewett, Frank B., Alexander Dallas Bache, a founder, first president and benefactor of the National Academy of Sciences, 181
Johnston, H. F., Magnetic work at sea, 257

Köhler, Wolfgang, On the nature of associations, 489

Kopal, Zdeněk, Determination of limb darkening in eclipsing binaries from color-index observations, 565
Korff, S. A., The production of neutrons by the cosmic radiation, 589
Krautheimer, Richard, S. Pietro in Vincoli and the tripartite transept in the Early Christian basilica, 353

Landis, Carney, Psychoanalysis and scientific method, 515
Lashley, Karl S., Coalescence of neurology and psychology, 461

McComb, H. E., Geomagnetic observatories and instruments, 239
McNish, A. G., The significance of fossil magnetism, 225
Magnetism, see geomagnetism

Natural selection before the "Origin of Species" (Zirkle), 71
Nervous system, central, Autonomous versus reflexogenous activity of (Weiss), 53
Neutrons, Production of, by the cosmic radiation (Korff), 589

Odgers, Merle M., Bache as an educator, 161
Olschki, Leonardo, What Columbus saw on landing in the West Indies, 633
"Origin of Species," Natural selection before (Zirkle), 71

Peeper, male spring, Reproductive physiology of (Rugh), 617
Psychology, Symposium on recent advances in, 461

Radio transmission across the Atlantic, short wave, Correlations of, with magnetic conditions (Hallborg), 323
Rugh, Roberts, Experimental studies on the reproductive physiology of the male spring peeper, *Hyla crucifer*, 617

S. Pietro in Vincoli and the tripartite transept in the Early Christian basilica (Krautheimer), 353
Slavs, John Bowring and the poetry of (Coleman), 431
String figures, Aboriginal Australian (Davidson), 763
Supernova spectra, Theoretical synthesis of (Whipple and Gaposchkin), 1
Symposia:
  Geomagnetism, 187
  Recent advances in psychology, 461

Thorndike, Edward L., Mental abilities, 503
Tolman, Edward C., Motivation, learning and adjustment, 543

Weiss, Paul, Autonomous versus reflexogenous activity of the central nervous system, 53
Whipple, Fred L., and Cecilia Payne-Gaposchkin, Theoretical synthesis of supernova spectra, 1

Yeast, Effects of heat and ultraviolet light on certain physiological properties of (Anderson and Duggar), 661
Yerkes, Robert M., Psychology and defense, 527

Zirkle, Conway, Natural selection before the "Origin of Species," 71

ISBN 1-4223-7222-7

9 781422 372227

Copyrighted material